

UNIVERSIDADE FEDERAL DO PARANÁ

ANDRÉ FENIAS MOIANE

INTEGRAÇÃO DO ALGORÍTMO DE CLUSTERIZAÇÃO PROPAGAÇÃO DE
AFINIDADE E O CLASSIFICADOR PARAMÉTRICO MÁXIMA
VEROSSIMILHANÇA PARA CLASSIFICAÇÃO DE IMAGENS HIPERSPECTRAIS

CURITIBA

2018

ANDRÉ FENIAS MOIANE

INTEGRAÇÃO DO ALGORÍTMO DE CLUSTERIZAÇÃO PROPAGAÇÃO DE
AFINIDADE E O CLASSIFICADOR PARAMÉTRICO MÁXIMA
VEROSSIMILHANÇA PARA CLASSIFICAÇÃO DE IMAGENS HIPERSPECTRAIS

Tese apresentada ao curso de Pós-Graduação em
Ciências Geodésicas, Setor de Ciências da Terra,
Universidade Federal do Paraná, como requisito
parcial à obtenção de grau de Doutor em Ciências
Geodésicas.

Orientador: Prof. Dr. Álvaro Muriel Lima Machado

CURITIBA

2018

FICHA CATALOGRÁFICA ELABORADA PELO SISTEMA DE BIBLIOTECAS/UFPR
BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

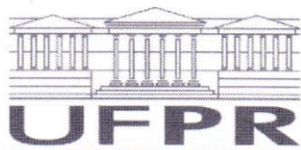
M712l Molane, André Fenias
Integração do algoritmo de clusterização propagação de afinidade e o classificador paramétrico
máxima verossimilhança para classificação de imagens hiperspectrais / André Fenias Molane. –
Curitiba, 2018.

Tese (Doutorado) - Universidade Federal do Paraná, Setor de Ciências da Terra, Programa de
Pós-Graduação em Ciências Geodésicas, 2018.

Orientador: Prof. Dr. Alvaro Muriel Lima Machado .

1. Máxima verossimilhança. 2. Imagens hiperspectrais. 3. Clusterização de dados.
I. Universidade Federal do Paraná. II. Machado , Alvaro Muriel Lima. III. Título.

CDD: 526.1



MINISTÉRIO DA EDUCAÇÃO
SETOR SETOR DE CIÊNCIAS DA TERRA
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO CIÊNCIAS
GEODÉSICAS

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em CIÊNCIAS GEODÉSICAS da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **ANDRE FENIAS MOIANE** intitulada: **INTEGRAÇÃO DO ALGORÍTMO DE CLUSTERIZAÇÃO PROPAGAÇÃO DE AFINIDADE E O CLASSIFICADOR PARAMÉTRICO MÁXIMA VEROSSIMILHANÇA PARA CLASSIFICAÇÃO DE IMAGENS HIPERSPECTRAIS**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 06 de Setembro de 2018.


ALVARO MURIEL LIMA MACHADO

Presidente da Banca Examinadora (UFPR)


JORGE ANTONIO SILVA CENTENO

Avaliador Interno (UFPR)


SELMA REGINA ARANHA RIBEIRO

Avaliador Externo (UEPG)


HIDEO ARAKI

Avaliador Interno (UFPR)

Dedico esta tese à memória de meus pais Fenias Moiane e Atália Silva Machavane, que me educaram na cultura do trabalho e na verticalidade do carácter.

Também dedico à minha esposa Trafina Jorge Bila, e aos meus filhos André Moiane Júnior, Neil André Moiane e Cármen André Moiane, pela paciência e apoio durante todo o tempo em que estive desenvolvendo esta tese.

AGRADECIMENTOS

A Deus eterno pelo dom da vida, pela bênção e proteção, e por nunca ter me abandonado, me fortalecendo sempre que precisei.

Ao Prof. Dr. Álvaro Muriel Lima Machado, meu orientador, agradeço pela paciência, compreensão, dedicação, pelos ensinamentos e pelo apoio incondicional, para que este trabalho se tornasse realidade. Agradeço também pelas sábias palavras de motivação, dizendo sempre “André, o observador faz a realidade, trabalhe e trabalhe duro, o doutorado é seu”.

Ao Prof. Dr. Jorge Antônio Silva Centeno, pelos conhecimentos transmitidos, pela simplicidade, acessibilidade, disponibilidade de tempo, material, apoio moral e especialmente por ter acreditado em mim, e ter me aceitado no programa vai o meu profundo agradecimento.

À Prof. Dra. Luciene Delazari, coordenadora do curso de pós-graduação em ciências geodésicas da UFPR, pela simplicidade, acessibilidade, disponibilidade de tempo, material e pela confiança e aceitação no programa.

Ao Prof. Dr. Hideo Araki, pelos conhecimentos transmitidos, pela simplicidade, acessibilidade e disponibilidade para esclarecimento de dúvidas sempre que precisei, vão os meus sinceros agradecimentos.

Ao Prof. Dr. Edson Aparecido Mitishita, pelos conhecimentos transmitidos, pela acessibilidade e disponibilidade para esclarecimento de dúvidas, vai o meu profundo agradecimento.

Aos meus professores que transmitiam não só conhecimentos, mas também relações humanas, pela disponibilidade para esclarecimento de duvidas, vai o meu reconhecimento em especial ao Prof. Dr. Wander da Cruz, Prof. Dr. Pedro Faggion, Prof. Dr. Silvio Freitas, Prof. Dr. Quintino Dalmolin e Prof. Dr. Leonardo Ercolin.

Agradeço às instituições que tornaram a minha formação possível, nomeadamente INFATEC, MINAG e MITADER pela concessão da bolsa de estudos; à Coordenação do Programa de Pós-Graduação em Ciências Geodésicas da UFPR, pela confiança e aceitação no programa; e vai um agradecimento especial ao ex Diretor Nacional dos Recursos Humanos do MINAG, Sr. Paulo Azevedo da Graça, pelo apoio e celeridade na tramitação do expediente referente a bolsa de estudos.

Aos meus colegas, companheiros dos choros e risadas da vida acadêmica, obrigado por me terem proporcionado momentos inesquecíveis e apoio no decurso

desta jornada, especialmente ao Miguel Muguio, Caisse Amisse, Francisco Soares de Lima, Josenilde Janguia, Felipe Costa, Mário Palma, Maria, João Victor, José, Carla Casaroti, Niarkios Graça, Marlo Antônio, Leonardo Biffi, Júlio César, Idalécio Joaquim, Leonel Muthemba, Verónica Wamusse, Suzana Daniela, Andrea Galudht, Leandro, Ilich Contreras, Evelin, Felipe Carvajal e Carolina Aguilar. VOCÊS SÃO OS CARAS.

À Mônica, secretária do PPGCG da UFPR, pela eficiência, paciência e atenção na tramitação de documentos sempre que solicitei, o meu muito obrigado.

Gostaria de reconhecer e expressar a minha sincera gratidão a todas as organizações e pessoas que direta ou indiretamente me deram os seus enormes apoios e contribuições para finalizar esta pesquisa.

Por último, mas não menos importante, à minha família e amigos, embora não existam palavras que possam expressar a minha imensa gratidão pela força e apoio dado nos momentos mais estressantes e solitários da minha vida. Os meus agradecimentos especiais vão à minha esposa Fina e meus filhos Júnior, Neil e Cármén, pelo seu amor incondicional, compreensão e apoio total enquanto eu trabalhava nesta tese.

“Education is not the learning of facts, it’s rather the training of the mind to think.”
Albert Einstein

“Discipline is the bridge between goals and accomplishment.”
Jim Rohn

“It always seems impossible until it’s done.”
Nelson Mandela

RESUMO

Esta tese investiga um método de classificação alternativo que integra o algoritmo de clusterização Propagação de Afinidade baseado nas Classes (PAC) e o Classificador Máxima Verossimilhança (MAXVER) com vista a superar as limitações do MAXVER na classificação de dados de alta dimensionalidade e, assim, melhorar a sua acurácia. O novo classificador foi designado PAC-MAXVER, e compreende duas abordagens, seleção de características espectrais e classificação de imagem. O algoritmo de clusterização PAC foi usado para realizar a redução de dimensionalidade da imagem e seleção de características enquanto o MAXVER foi utilizado para a classificação da imagem. O desempenho do MAXVER em termos de acurácia da classificação e tempo de processamento é determinado em função da taxa de seleção realizada na fase de clusterização PAC. O desempenho de PAC-MAXVER foi avaliado e validado usando duas cenas hiperespectrais do AVIRIS (Airborne Visible Infrared Imaging Spectrometer) e HYDICE (Hyperspectral Digital Imagery Collection Experiment). Os resultados da classificação mostram que PAC-MAXVER observou uma enorme melhoria na acurácia, atingindo 94,15% e 96,47%, respectivamente para imagem AVIRIS e HYDICE se comparado com o MAXVER, que obteve 85,42% e 81,50%. Esses valores obtidos pelo PAC-MAXVER melhoraram a acurácia da classificação MAXVER em 8,73% e 14,97% para essas imagens. Os resultados também mostraram que o PAC-MAXVER teve um bom desempenho, mesmo para as classes com número limitado de amostras de treinamento, superando as limitações do MAXVER.

Palavras-chave: Propagação de Afinidade. Máxima Verossimilhança. Seleção de bandas. Integração. Classificação. Imagens hiperespectrais.

ABSTRACT

This thesis investigates an alternative classification method that integrates Class-based Affinity Propagation (CAP) clustering algorithm and Maximum Likelihood Classifier (MLC) with the purpose of overcome the MLC limitations in the classification of high dimensionality data, and thus improve its accuracy. The new classifier was named CAP-MLC, and comprises two approaches, spectral feature selection and image classification. CAP clustering algorithm was used to perform the image dimensionality reduction and feature selection while the MLC was employed for image classification. The performance of MLC in terms of classification accuracy and processing time is determined as a function of the selection rate achieved in the CAP clustering stage. The performance of CAP-MLC has been evaluated and validated using two hyperspectral scenes from the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) and the Hyperspectral Digital Imagery Collection Experiment (HYDICE). Classification results show that CAP-MLC observed an enormous improvement in accuracy, reaching 94,15% and 96,47% respectively for AVIRIS and HYDICE images if compared with MLC, which had 85,42% and 81,50%. These values obtained by CAP-MLC improved the MLC classification accuracy in 8,73% and 14,97% for these images. The results also show that CAP-MLC performed well, even for classes with limited training samples, surpassing the limitations of MLC.

Keywords: Affinity Propagation. Maximum Likelihood Classifier. Band selection. Integration. Classification. Hyperspectral imagens.

LISTA DE ILUSTRAÇÕES

FIGURA 1 - ESTRUTURA BÁSICA DE DADOS DE UMA IMAGEM HIPERESPECTRAL.....	24
FIGURA 2 - ESQUEMA DO MÉTODO DE CLASSIFICAÇÃO DISTÂNCIA MÍNIMA.	34
FIGURA 3 - DUAS AMOSTRAS ESPECTRAIS COM MESMO CENTROIDE E DISPERSÕES DIFERENTES.	35
FIGURA 4 - ESQUEMA DO MÉTODO DE CLASSIFICAÇÃO DA DISTÂNCIA DE MAHALANOBIS.	36
FIGURA 5 - GRÁFICO DE FATORES DA PROPAGAÇÃO DE AFINIDADE.	38
FIGURA 6 - DUAS MENSAGENS SÃO TROCADAS ENTRE PONTOS: (A) “RESPONSABILIDADES” $R(I,K)$ SÃO ENVIADAS DO PONTO I PARA O EXEMPLAR CANDIDATO K, E (B) “DISPONIBILIDADES” $A(I,K)$ SÃO ENVIADAS DO EXEMPLAR CANDIDATO K PARA O PONTO I.	39
FIGURA 7 - EXEMPLO DO EFEITO DE ATENUAÇÃO DE RUÍDO EM UMA IMAGEM DE RADAR, (A) ANTES E (B) DEPOIS DO USO DO FILTRO DA MEDIANA.	45
FIGURA 8 - IMAGEM AVIRIS DO LOCAL DE ESTUDO E VERDADE DE CAMPO.	51
FIGURA 9 - IMAGEM HYDICE DA ÁREA DE ESTUDO E VERDADE DE CAMPO.	53
FIGURA 10 - FLUXOGRAMA DA IMPLEMENTAÇÃO DA METODOLOGIA PROPOSTA PARA O MÉTODO PAC - MAXVER.....	54
FIGURA 11 - FLUXOGRAMA DA IMPLEMENTAÇÃO DA METODOLOGIA PROPOSTA PARA O MÉTODO PA - MAXVER.	54
FIGURA 12 - ESTRUTURA DE UMA IMAGEM HIPERESPECTRAL	57
FIGURA 13 - REPRESENTAÇÃO DO CUBO DA IMAGEM HIPERESPECTRAL E CADA IMAGEM BANDA COMO PONTO.....	58
FIGURA 14 - IMAGENS CLASSIFICADAS SEM APLICAÇÃO DO FILTRO DE MEDIANA (A), COM APLICAÇÃO DO FILTRO DE MEDIANA (B) E VERDADE DE CAMPO PARA O EXPERIMENTO I.	66
FIGURA 15 - IMAGENS CLASSIFICADAS SEM APLICAÇÃO DO FILTRO DA MEDIANA (A), COM APLICAÇÃO DO FILTRO DA MEDIANA (B) E VERDADE DE CAMPO PARA O EXPERIMENTO II.	68

FIGURA 16 - IMAGEM CLASSIFICADA USANDO O EXPERIMENTO III E VERDADE DE CAMPO.....	69
FIGURA 17 - IMAGEM CLASSIFICADA USANDO O EXPERIMENTO IV E VERDADE DE CAMPO.....	71
FIGURA 18 - COMPARAÇÃO DAS CLASSIFICAÇÕES NOS EXPERIMENTOS I – IV.....	72
FIGURA 19 - IMAGEM CLASSIFICADA COM APLICAÇÃO DO FILTRO DA MEDIANA, IMAGEM ORIGINAL E VERDADE DE CAMPO.....	75
FIGURA 20 - IMAGEM CLASSIFICADA COM APLICAÇÃO DO FILTRO DA MEDIANA PARA O EXPERIMENTO II, IMAGEM ORIGINAL E VERDADE DE CAMPO.....	77
FIGURA 21 - IMAGEM CLASSIFICADA PARA O EXPERIMENTO III, IMAGEM ORIGINAL E VERDADE DE CAMPO.	78
FIGURA 22 - IMAGEM CLASSIFICADA PARA O EXPERIMENTO IV, IMAGEM ORIGINAL E VERDADE DE CAMPO.	79
FIGURA 23 - COMPARAÇÃO DAS CLASSIFICAÇÕES NOS EXPERIMENTOS I – IV.....	80
FIGURA 24 - COMPARAÇÃO DAS ACURÁCIAS DAS CLASSIFICAÇÕES NOS EXPERIMENTOS I – IV PARA AS IMAGENS AVIRIS E HYDICE.....	83

LISTA DE TABELAS

TABELA 1 - INTERPRETAÇÃO DOS VALORES DO COEFICIENTE KAPPA (CK).	49
TABELA 2 - CLASSES E VERDADES DE CAMPO PARA TREINAMENTO E TESTE PARA AVIRIS.....	52
TABELA 3 - CLASSES E VERDADES DE CAMPO PARA TREINAMENTO E TESTE PARA HYDICE.....	53
TABELA 4 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO I COM A IMAGEM AVIRIS.	66
TABELA 5 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO II COM A IMAGEM AVIRIS.	68
TABELA 6 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO III COM A IMAGEM AVIRIS.	70
TABELA 7 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO IV COM A IMAGEM AVIRIS.	71
TABELA 8 - RESUMO DA ACURÁCIA DOS EXPERIMENTOS I – IV COM IMAGEM AVIRIS.	73
TABELA 9 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO I COM A IMAGEM HYYDICE.....	76
TABELA 10 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO II COM A IMAGEM HYDICE.	77
TABELA 11 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO III COM A IMAGEM HYDICE.	78
TABELA 12 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO IV COM A IMAGEM HYDICE.	79
TABELA 13 - RESUMO DA ACURÁCIA DOS EXPERIMENTOS I – IV COM IMAGEM HYDICE.	81

LISTA DE ABREVIATURAS E SIGLAS

AG	-Acurácia Global
APr	-Acurácia do Produtor
AU	-Acurácia do Usuário
AVIRIS	-Airborne Visible Infrared Imaging Spectrometer
CK	-Coeficiente Kappa
CPU	-Unidade de Processamento Central
DM	-Distância Mínima
EI	-Erros de Inclusão
ENVI	-Ambiente para Visualização de Imagens
EO	-Erros de Omissão
HYDICE	-Experimento para Coleção de Imagem Digital Hiperspectral
ID	-Divergência de Informação
ISODATA	-Iterative Self-Organizing Data Analysis Technique
JM	-Jeffries-Matusita
LDA	-Análise Discriminante Linear
MATLAB	-Laboratório de Matriz
MAXVER	-Máxima Verossimilhança
MVPCA	-PCA de Variância Máxima
PAC	-Propagação de Afinidade baseada nas Classes
PA	-Propagação de Afinidade
PCA	-Análise de Componentes Principais
RAM	-Memória de Acesso Aleatório
RGB	-Red, Green and Blue
SAM	-Mapeador de Ângulo Espectral
SID	-Divergência de Informação Espectral
SS	-Soma Quadrática
SVM	-Máquina de Vetor de Suporte

SUMÁRIO

CAPÍTULO 1. INTRODUÇÃO	16
1.1 VISÃO GERAL	16
1.2 OBJETIVOS DA TESE.....	19
1.2.1 Objetivo Geral	19
1.2.2 Objetivos Específicos.....	19
1.3 JUSTIFICATIVA	20
1.4 PROBLEMA DE PESQUISA.....	21
1.5 HIPÓTESE	21
1.6 CONTRIBUIÇÕES DA TESE	22
1.6.1 Artigos Produzidos Durante a Tese	23
1.7 ORGANIZAÇÃO DA TESE	23
CAPÍTULO 2. REVISÃO DE LITERATURA	24
2.1 IMAGENS HIPERESPECTRAIS	24
2.2 MÉTODOS DE CLASSIFICAÇÃO DE IMAGENS HIPERESPECTRAIS	25
2.2.1 Métodos de Classificação não Supervisionados	26
2.2.2 Métodos de Classificação Supervisionados.....	29
2.3 MÉTODOS DE REDUÇÃO DE DIMENSIONALIDADE DE IMAGENS HIPERESPECTRAIS	37
2.3.1 Propagação de Afinidade.....	38
2.3.2 Análise de Componentes Principais	41
2.4 FILTRAGENS DE IMAGENS	44
2.4.1 Filtragem pela Mediana	44
2.5 AVALIAÇÃO DA ACURÁCIA E VALIDAÇÃO DA CLASSIFICAÇÃO	46
CAPÍTULO 3. MATERIAL E MÉTODOS	50
3.1 MATERIAL	50
3.2 LOCALIZAÇÃO DA ÁREA DE ESTUDO E IMAGENS UTILIZADAS	50
3.2.1 Imagem AVIRIS	51
3.2.2 Imagem HYDICE	52
3.3 MÉTODOS	53
3.3.1 Pré-processamento das imagens	55
3.3.2 Processamento	55

3.3.3 Classificação da imagem	63
3.3.4 Avaliação da Acurácia e Validação dos Resultados	64
3.3.5 Pós – processamento da Imagem Classificada	65
CAPÍTULO 4. APRESENTAÇÃO E ANÁLISE DOS RESULTADOS.....	66
4.1 EXPERIMENTOS COM A IMAGEM AVIRIS.....	66
4.1.1 Experimento I – Método PAC – MAXVER	66
4.1.2 Experimento II – Método PA – MAXVER	68
4.1.3 Experimento III – Método PCA – MAXVER	69
4.1.4 Experimento IV – Método MAXVER	71
4.1.5 Análise Conjunta dos Resultados Obtidos pelos Métodos I – IV, Usando a Imagem Hiperespectral AVIRIS	72
4.2 EXPERIMENTOS COM A IMAGEM HYDICE	75
4.2.1 Experimento I – Método PAC – MAXVER	75
4.2.2 Experimento II – Método PA – MAXVER	77
4.2.3 Experimento III – Método PCA – MAXVER	78
4.2.4 Experimento IV – Método – MAXVER	79
4.2.5 Análise Conjunta dos Resultados Obtidos pelos Métodos I – IV, Usando a Imagem Hiperespectral HYDICE	80
4.3 ANÁLISE CONJUNTA DOS RESULTADOS OBTIDOS NAS DUAS IMAGENS	83
CAPÍTULO 5. CONCLUSÕES E TRABALHOS FUTUROS.....	85
5.1 CONCLUSÕES	85
5.2 TRABALHOS FUTUROS	87
REFERÊNCIAS.....	89

CAPÍTULO 1. INTRODUÇÃO

1.1 VISÃO GERAL

O Sensoriamento remoto é um processo que envolve a coleta e interpretação de informação referente a um objeto, área ou evento sem que haja contato físico com os mesmos. Nesse processo, aeronaves e satélites são as plataformas comuns usadas para sensoriar a terra e seus recursos naturais (GOETZ et al., 1985). O uso de fotografias aéreas na faixa do visível do espectro eletromagnético é a forma mais antiga de sensoriamento remoto, mas os desenvolvimentos tecnológicos permitem a aquisição de informação em outros comprimentos de onda incluindo o infravermelho próximo, infravermelho termal e microondas (ABLIN; SULOCHANA, 2013).

A medição de energia nas mais estreitas e numerosas bandas em até 200 (ou mais) bandas espectrais contíguas é denominada imagem hiperespectral (ABLIN; SULOCHANA, 2013). As inúmeras bandas estreitas dos sensores hiperespectrais fornecem medição espectral contínua em todo o espectro eletromagnético e, portanto, são mais sensíveis a variações sutis de energia refletida (CAMPBELL, 2002).

As imagens hiperespectrais contêm atributos espectrais extremamente ricos que fornecem dados mais detalhados sobre as propriedades espectrais dos materiais, permitindo uma discriminação mais precisa dos objetos se comparadas com as tradicionais imagens *RGB* ou mesmo multispectrais (LEE; LANDGREBE, 1993; ABLIN; SULOCHANA, 2013; CHEHDI; SOLTANI; CARIOU, 2014). Com esses dados, torna-se possível fornecer grande quantidade de informação espectral e espacial para aplicações em análise de imagens, mapeamento ao nível do subpixel e detecção de alvos (CHANG; LANDGREBE, 2003, 2002). Desta forma, a análise de imagens hiperespectrais é de interesse crescente em aplicações tais como planejamento urbano, agricultura, estudo e monitoramento de florestas, detecção de sutis mudanças na vegetação (COSTA et al., 2007; SAHOO et al., 2007; PEDDLE et al., 2008), solo, água e reflectância mineral (LEVESQUE; STAENZ, 2008).

Apesar do enorme potencial em discriminar objetos, a classificação e análise dessas imagens ainda constitui tema de pesquisa importante e desafiador no reconhecimento de padrões e em estudos de sensoriamento remoto (BARTELS; WEI, 2006; BRZANK; HEIPKE, 2007).

Vários métodos foram propostos para classificação de imagens hiperespectrais. Roessner et al. (2001), combinaram a classificação por Máxima Verossimilhança (MAXVER) com separação linear de mistura espectral incluindo um novo modelo matemático para seleção de materiais com propriedades espectrais constantes (*endmembers*) orientado ao pixel para acomodar o grande número desses materiais ocorrendo no ambiente urbano. Outras abordagens de classificação de dados hiperespectrais são baseadas na MAXVER (JIA, 2002), Árvores de Decisão (GOEL et al., 2003), Redes Neurais (DEL FRATE et al., 2007), Algoritmos Genéticos (VAIPHASA, 2003) e técnicas baseadas em Kernel (MULLER et al., 2001; CAMPS-VALLS; BRUZZONE, 2005).

Um dos classificadores paramétricos padrões comumente usado é a Máxima Verossimilhança (RICHARDS, 2013). Este é um método supervisionado de classificação de imagens, e usa amostras de treinamento de grande tamanho se comparado aos métodos não paramétricos (PENG et al., 2016). Uma questão relacionada aos classificadores paramétricos em geral, e particularmente ao Máxima Verossimilhança é que a acurácia da classificação obtida é tipicamente comprometida quando aplicada a imagens hiperespectrais para classes com pequenas amostras de treinamento, produzindo resultados de classificação inconsistentes (MARANGOZ et al., 2006; CAMPS-VALLS et al., 2014). Enquanto isso, classificadores discriminativos tais como Propagação de Afinidade, Máquina de Vetor de Suporte e outros procuram descrever totalmente uma classe pela modelagem da função de densidade de probabilidade (ANDREW; MICHAEL, 2001). Estes classificadores precisam apenas de pequenas amostras de treinamento, embora estejam sujeitas a falta de robustez no que respeita a ruídos e super ajustamento (PENG et al., 2016).

A Propagação de Afinidade (PA) é um algoritmo de *clusterização* baseada em exemplares ou centros de *clusters*. Este algoritmo usa como dados de entrada as similaridades entre pares de pontos ou de imagens, e a partir da propagação de mensagens de afinidade entre pontos num gráfico de fatores, produz um conjunto de exemplares que melhor representam os pontos (FREY; DUECK, 2007).

Após identificação de exemplares, os pontos ou imagens não identificados são atribuídos ao exemplar mais apropriado. Comparado com as abordagens tradicionais, a PA apresenta enorme potencial na *clusterização* de dados não

métricos e adequados para análise exploratória de dados, usando medidas de similaridade não comuns (GUAN et al., 2011).

Recentemente, PA foi aplicada na classificação de dados de sensoriamento remoto hiperespectral. Jia, Qian e Ji (2008), Qian, Yao e Jia (2009), Jia et al. (2012) propuseram a propagação de afinidade para seleção de bandas em imagens hiperespectrais. Os resultados experimentais em duas disposições de dados mostraram que o método proposto representa melhor a imagem hiperespectral, do ponto de vista da classificação da imagem por pixel, se comparados com alguns métodos recentes tais como ID (Information Divergence), MVPCA (Maximum Variance PCA) e K-centers.

Chehdi, Soltani e Cariou (2014), propuseram a classificação de pixels em imagens hiperespectrais usando a Propagação de Afinidade. O método consiste na redução do número de pixels a serem classificados e aplicar PA por partição de blocos. Os resultados experimentais em várias imagens mostraram a eficiência do método. Yang et al. (2010) propuseram Propagação de Afinidade baseada em estatística fuzzy para classificação de imagens multispectrais. Segundo os autores, os resultados experimentais demonstraram boa eficiência computacional e alta acurácia ao comparar com o *K-means* e *fuzzy K-means*. Yang et al. (2013) propuseram Propagação de Afinidade incremental e decremental para classificação semi-supervisionada de imagens multispectrais e afirmam terem obtido bons resultados.

Considerando o grande potencial do *clusterizador* Propagação de Afinidade e do classificador Máxima Verossimilhança, esta tese propõe sua integração visando melhorar a acurácia da classificação de imagens hiperespectrais, mesmo ao aplicar sobre pequenas amostras de treinamento. O grande diferencial entre esta tese e as várias abordagens apresentadas na literatura reside nos seguintes aspectos:

- 1) Propõe integrar PA e MAXVER, de modo a explorar suas vantagens de forma combinada;
- 2) Propõe selecionar bandas das imagens hiperespectrais mediante *clusterização* supervisionada baseada nas classes;
- 3) Propõe combinar as bandas selecionadas e aplicar nova *clusterização* não supervisionada, baseada nas bandas, para obtenção das bandas finais usadas na classificação da imagem.

1.2 OBJETIVOS DA TESE

1.2.1 Objetivo Geral

Esta tese tem como objetivo principal desenvolver uma abordagem efetiva que integra os conceitos de *clusterização* por Propagação de Afinidade baseada nas Classes e a classificação por Máxima Verossimilhança, combinando suas vantagens em forma sinérgica, minimizando as limitações associadas à aplicação separada de cada uma dessas técnicas.

1.2.2 Objetivos Específicos

1. Extrair/selecionar características em imagens hiperespectrais, usando PA, avaliando suas vantagens e desvantagens na classificação e interpretação de imagens.
2. Classificar tipos de coberturas de terra em imagens hiperespectrais, mediante MAXVER, avaliando seu potencial na classificação e interpretação de imagens.
3. Desenvolver uma nova abordagem para classificação de imagens hiperespectrais, baseada na integração da técnica de *clusterização* Propagação de Afinidade baseado na Classe (PAC) com o classificador paramétrico Máxima Verossimilhança, e analisar as vantagens que podem ser obtidas a partir da exploração conjunta dessas abordagens.
4. Avaliar a nova técnica de classificação desenvolvida neste trabalho, resultante da integração dos conceitos de *clusterização* e de classificação.
5. Efetuar um estudo comparativo entre a abordagem desenvolvida e outras técnicas tais como Análise de Componentes Principais, a Propagação de Afinidade original e a Máxima Verossimilhança, utilizando dados hiperespectrais reais obtidos pelo sensor AVIRIS (Airborne Visible Infra-Red Imaging Spectrometer) e HYDICE (Hyperspectral Digital Imagery Collection Experiment).

1.3 JUSTIFICATIVA

Dados em alta dimensionalidade podem oferecer um poder discriminante bem mais elevado do que dados tradicionais em baixa dimensionalidade (LEE; LANDGREBE, 1993; JIMENEZ; LANDGREBE, 1999; SERPICO; BRUZZONE, 2001; ABLIN; SULOCHANA, 2013). Fukunaga (1990) demonstra que classes espectralmente muito semelhantes entre si (classes que compartilham vetores de médias muito próximas) podem ser separadas satisfatoriamente em espaços de dimensão mais alta. Esta é uma das motivações para o desenvolvimento de sistemas sensores com um número grande de bandas espectrais, conhecidos como sensores hiperespectrais.

Entretanto, uma das principais dificuldades que surgem no processo de classificação de dados em alta dimensionalidade por meio de classificadores paramétricos diz respeito ao número de amostras de treinamento (em geral limitado) em comparação com o número de parâmetros a serem estimados (CAMPS-VALLS et al., 2014). Um número limitado de amostras de treinamento resulta em uma estimativa pouco confiável dos parâmetros em um classificador paramétrico como, por exemplo, a Máxima Verossimilhança (MAXVER) e, conseqüentemente em uma acurácia menor da imagem temática produzida (PENG et al., 2016).

Iniciando o processo de classificação com dados em dimensionalidade reduzida, a acurácia da imagem temática tende, inicialmente, a aumentar na medida em que informações (na forma de bandas espectrais) adicionais são incluídas. Em um determinado momento, a acurácia atinge um máximo para em seguida passar a diminuir, na medida em que a dimensionalidade dos dados continua a aumentar. Este problema, conhecido pela comunidade internacional como *fenômeno de Hughes*, vem sendo objeto de estudo por pesquisadores como HOFFBECK e LANDGREBE (1996) e JIMENEZ e LANDGREBE (1999), por exemplo. Redução na dimensionalidade dos dados por meio de técnicas de extração ou seleção de variáveis, introdução de amostras de treinamento semi-rotuladas, técnicas de análise discriminante regularizada, são abordagens que vem sendo investigadas com o objetivo de minimizar as consequências de tal fenômeno. Neste contexto, desperta o interesse a utilização de classificadores não paramétricos, como é o caso da propagação de afinidade, que apresenta a vantagem de não ser afetado por este tipo de problema (ANDREW; MICHAEL, 2001; HUANG; DAVIS; TOWNSHEND,

2002). Estes classificadores, precisam somente de uma pequena amostra de treinamento embora sofram de falta de robustez no que respeita a ruídos e super ajustamento (PENG et al., 2016). No contexto da classificação de imagens hiperespectrais, algumas investigações experimentais apontam a eficácia da PA para a análise desses dados.

Nesta tese investiga-se a integração do algoritmo de *clusterização* PAC com o classificador paramétrico MAXVER na classificação de imagens hiperespectrais de sensoriamento remoto, nas quais várias classes estão presentes. Na metodologia proposta, investiga-se a implementação da PAC como etapa de préprocessamento para o classificador paramétrico acima mencionado. A seleção de bandas por PA permite classificar na imagem hiperespectral até aquelas classes com pequenas amostras de treinamento, contornando, desta forma, a principal limitação apresentada pelo classificador MAXVER. Uma vantagem adicional da PA reside na combinação da classificação baseada em modelos com a baseada na afinidade, o que lhe confere um maior poder discriminante entre pares de bandas próximos para verificar suas similaridades (SHENTAL et al., 2003; ROSALES; FREY, 2003), e descreve os dados como uma mistura de distribuição condicional de classes (FREY; DUECK, 2006).

1.4 PROBLEMA DE PESQUISA

Em que medida a integração do algoritmo de *clusterização* Propagação de Afinidade baseado nas Classes com o classificador paramétrico Máxima Verossimilhança melhora a acurácia da classificação em imagens hiperespectrais?

1.5 HIPÓTESE

Peng et al. (2016), Bandos et al. (2009), Li e Qian (2011) asseveram que classificadores generativos tais como MAXVER, LDA e outros que modelam a distribuição de probabilidade conjunta $p(x, y)$, das entradas x e da etiqueta y , e fazem suas previsões usando regras de Bayes para calcular a probabilidade *a posteriori* ou condicional $p(y | x)$, não são sensíveis a ruídos, mas precisam de grande quantidade de dados de treinamento para alcançar boa precisão na

discriminação e classificação de feições. Enquanto isso, os classificadores discriminativos como a PA, SVM, etc, modelam a distribuição de probabilidade *a posteriori* ou condicional $p(y | x)$ diretamente, ou aprendem um mapa direto a partir das entradas x para as etiquetas y da classe, precisam somente de pequena amostra para treinamento, mas são muito sensíveis a ruídos na imagem e ao super ajustamento.

Shental et al. (2003), Rosales e Frey (2003), asseveram que PA combina as vantagens da classificação baseada em modelos e da baseada na afinidade. Diferentemente das técnicas que constroem e aprendem modelos de probabilidade de transições entre pontos, PA por si aprende um modelo de probabilidade de dados. Frey e Dueck (2006), afirmam que PA examina diretamente pares próximos de pontos para ajudar a verificar se podem ou não pertencer ao mesmo *cluster*. E, como classificador baseado no modelo, a PA usa um modelo de probabilidade que descreve os dados como uma mistura de distribuição condicional de classes. Outra questão é que a acurácia obtida com os classificadores paramétricos padrões, comumente empregados, como é o caso da MAXVER, é tipicamente comprometida quando aplicada a imagens hiperespectrais (CAMPS-VALLS et al., 2014), produzindo resultados de classificação inconsistentes (MARANGOZ et al., 2006). Os métodos discriminativos tais como PA têm ganhado popularidade devido à sua habilidade em produzir resultados de classificação desejáveis (JIA; QIAN; JI, 2008; MEILA; SHI, 2001; NG; JORDAN; WEISS, 2002; STREHL; GHOSH, 2002; YANG et al., 2013; YANG et al., 2010. Com base nos fundamentos apresentados acima e demais literatura, é levantada a seguinte hipótese:

Hipótese: A integração do algoritmo de *clusterização* Propagação de Afinidade baseado nas Classes e o classificador Máxima Verossimilhança permite classificar imagens hiperespectrais utilizando pequenas amostras de treinamento e melhora a acurácia da classificação.

1.6 CONTRIBUIÇÕES DA TESE

Esta tese tem como contribuições o seguinte: desenvolvimento de um método alternativo para classificação de imagens hiperespectrais de alta dimensionalidade com uso limitado de amostras de treinamento. O método proposto permite que a

redução da dimensionalidade da imagem seja realizada numa primeira fase pela *clusterização* baseada em cada uma das classes existentes na imagem, e seguidamente pela *clusterização* conjunta de todas as bandas obtidas durante o processo de *clusterização* classe a classe. Este procedimento removeu somente as bandas altamente correlacionadas e com informação redundante e ruidosa, garantindo a melhora da acurácia da classificação, mantendo o detalhe nas feições.

1.6.1 Artigos Produzidos Durante a Tese

MOIANE, A. F.; MACHADO, A. M. L. Evaluation of the Clustering Performance of Affinity Propagation Algorithm considering the influence of Preference parameter and Damping factor. **Bulletin of Geodetic Sciences**, v. 24, n. 4, Brasil, 2018.

MOIANE, A. F.; MACHADO, A. M. L. Class-based Affinity Propagation for Hyperspectral image Dimensionality Reduction and Improvement of Maximum Likelihood Classification Accuracy. **Bulletin of Geodetic Sciences**, Brasil, 2018.

MOIANE, A. F.; MACHADO, A. M. L. Assessing the classification accuracy of AVIRIS hyperspectral image by varying affinity propagation preference parameter and damping factor. **Brazilian Journal of Cartography**, Brasil, 2018.

1.7 ORGANIZAÇÃO DA TESE

Os capítulos seguintes desta tese estão estruturados da forma seguinte:

No capítulo 2, são revistos conceitos sobre a tecnologia de imagens hiperespectrais, extração de características no sensoriamento remoto, métodos de classificação de imagens hiperespectrais, métodos de redução de dimensionalidade por PA e por PCA, filtragem de imagens pela mediana, e o capítulo termina com a apresentação dos métodos de avaliação da acurácia e validação da classificação.

O capítulo 3 introduz os materiais e o método proposto para classificação de feições em imagens hiperespectrais.

Os resultados obtidos pela implementação da metodologia proposta nesta tese são apresentados e discutidos no capítulo 4.

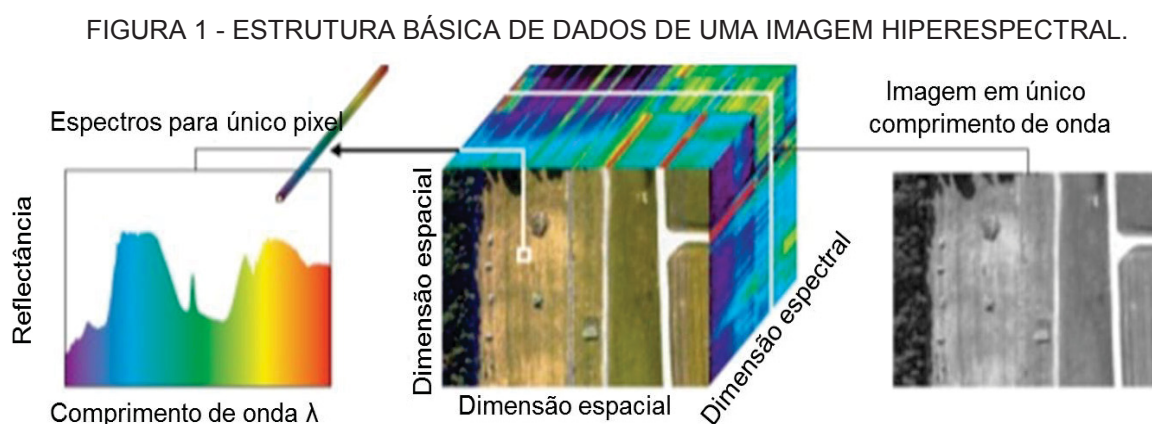
Para terminar, as conclusões do trabalho desenvolvido nesta e recomendações sobre pesquisas futuras referentes a classificação de imagens hiperespectrais são apresentados no capítulo 5.

CAPÍTULO 2. REVISÃO DE LITERATURA

2.1 IMAGENS HIPERESPECTRAIS

O sensoriamento remoto hiperespectral é uma das maiores adições recentes para a tecnologia de sensoriamento remoto. Este permite a aquisição simultânea de dados em centenas de bandas espectrais com larguras de banda estreitas e podem prover assinaturas espectrais contíguas detalhadas que os tradicionais sensores multispectrais não podem oferecer (LONG et al., 2015; CHUTIA et al., 2016).

As imagens hiperespectrais são melhor descritas como cubos ou paralelepípedos de imagens com duas dimensões espaciais e a terceira como dimensão espectral, como ilustrado na Figuras 1. Elas são usadas em uma variedade de domínios, incluindo geologia planetária e terrestre, agricultura, estudo de florestas, estudo de águas costeiras e internas, estudos atmosféricos, avaliação de desastres ambientais e estudos urbanos (MANOLAKIS; MARDEN; SHAW, 2003).



FONTE: MANOLAKIS; MARDEN; SHAW, 2003.

Na Figura 1 é ilustrado o caráter simultâneo (espectral e espacial) dos dados hiperespectrais. À esquerda, o cubo pode ser visualizado como uma disposição de espectros, um para cada pixel, ou como uma combinação de imagens, cada uma representa uma única banda espectral, conforme figura a direita.

Contudo, a maior dificuldade encontrada no sensoriamento remoto hiperespectral é a enorme dimensionalidade dos dados que faz com que surja o fenómeno de Hughes (HSU, 2007). Este fenómeno refere-se ao fato do tamanho da amostra necessária para treinar um classificador específico crescer exponencialmente com o número de bandas espectrais. Deve ser dada melhor

consideração à coleta de amostras de treinamento e de teste que caracterizam o intervalo de variabilidade da superfície da terra na escala espacial da imagem (PAL; MATHER, 2006). Entretanto, a largura de banda estreita contígua dos dados hiperespectrais facilita o levantamento detalhado das características da superfície da Terra, o que não seria possível com larguras de banda relativamente larga, adquiridas pelos sensores multispectrais. Consequentemente, sua capacidade em distinguir diferentes objetos no terreno e os descrever com detalhe é muito melhorada (WEN; YANG; HU, 2009).

O aumento na dimensionalidade espectral não apenas permite a definição da superfície da terra aos níveis de uso e cobertura de terra, mas também facilita potencialmente a caracterização de vários minerais (LEVESQUE; STAENZ, 2008), solos, rochas (GOLUBIEWSKI; WESSMAN, 2010), vegetação (PEDDLE et al., 2008; ZHANG et al., 2008; COSTA et al., 2007), classificação de uso de terra e inferência de propriedades químicas de áreas de vegetação heterogêneas (SAHOO et al., 2007). Contudo, o maior detalhe espectral vem a um custo de enormes volumes de dados com alta dimensionalidade, o que coloca grandes desafios em extrair informação temática a partir dessas imagens (ZHANG; QIU, 2012). Detecção de anomalias é um dos aspectos mais importantes na análise de imagens hiperespectrais, pois estas imagens agora podem descobrir muitas substâncias materiais, que anteriormente não foram resolvidas pelos sensores multispectrais (CHANG; CHIANG, 2002).

2.2 MÉTODOS DE CLASSIFICAÇÃO DE IMAGENS HIPERESPECTRAIS

Os métodos de classificação podem ser agrupados em duas categorias: métodos supervisionados e não supervisionados (DUDA; HART; STORK, 2000). Nos métodos de classificação supervisionada, as classes são previamente definidas pelo analista, ou seja, definidas ou caracterizadas pelas amostras de treinamento. Cada classe pode então ser caracterizada por uma função de decisão que pode ser de natureza probabilística (paramétrica) ou de natureza determinística (não paramétrica). Por outro lado, os métodos não supervisionados oferecem outra abordagem, em que a natureza (ou definição) de classes, e até mesmo o número de classes existentes são desconhecidos.

Essas duas categorias de métodos supervisionados e não supervisionados podem ainda ser subdivididas em classificadores generativos e discriminativos (NG; JORDAN, 2002). A principal diferença entre estes dois grupos é a seguinte: os classificadores generativos modelam a distribuição de probabilidade conjunta $p(x, y)$, das entradas x e do rótulo y , e fazem suas previsões usando regras de Bayes para calcular a probabilidade *a posteriori* ou condicional $p(y | x)$ e, em seguida, escolhem a classe y mais provável. Diferentemente dos métodos generativos, os discriminativos modelam a distribuição de probabilidade *a posteriori* ou condicional $p(y | x)$ diretamente, ou aprendem um mapa direto a partir das entradas x para os rótulos da classe.

2.2.1 Métodos de Classificação não Supervisionados

Nos classificadores não-supervisionados é o próprio classificador que identifica as amostras espectrais distintas que poderão existir na imagem, num processo iterativo, sendo os mais utilizados o K-MEANS e o ISODATA, também chamados algoritmos de agrupamento. Neste tipo de classificadores o objetivo é agrupar amostras espectrais semelhantes. Essas amostras espectrais não são conhecidas no início, podendo o número de amostras espectrais ser um parâmetro configurável. Depois de efetuada a classificação o usuário terá de inferir sobre a natureza das classes que o algoritmo classificou, isto é, fazer corresponder a cada classe identificada pelo algoritmo um tema que caracterize uma classe presente na imagem, e agrupar, se for o caso, algumas classes classificadas numa só, se verificar-se que poderão fazer parte de um tema único.

De uma maneira geral, estes classificadores começam por calcular as médias de amostras iniciais de pixels que se destacam pelo fato de se apresentarem uniformemente distribuídas na imagem. Em seguida associam cada pixel da imagem à amostra espectralmente mais próxima, criando assim um novo conjunto de amostras. Depois, com as novas amostras, recalculam as respectivas médias. Estes dois últimos passos são repetidos de modo iterativo até as distâncias entre os pixels e as médias das amostras serem menores do que um limite determinado pelo usuário.

No caso destas classificações, quanto maior a heterogeneidade das amostras, maior a certeza de que todas as classes possíveis estarão representadas.

A principal vantagem do uso destes tipos de classificação é o fato de não ser necessário um conhecimento prévio da área de estudo, enquanto a sua desvantagem é o fato de o utilizador ter pouco controle sobre a separação entre classes.

2.2.1.1 Método ISODATA

O método de *clusterização* ISODATA usa a distância espectral num processo feito em três etapas: (i) classifica iterativamente os pixels; (ii) com isso vai redefinindo os critérios para cada classe; (iii) para reclassificar a imagem até que os padrões de distância espectral nos dados apareçam. ISODATA significa Iterative Self-Organizing Data Analysis Technique. Ele é iterativo, uma vez que realiza repetidamente uma classificação inteira (produzindo uma imagem temática de saída) e recalcula as estatísticas.

Self-Organizing refere-se ao modo como o método localiza os agrupamentos com mínima informação fornecida pelo usuário. O classificador ISODATA fundamenta-se no princípio de que todos os agrupamentos encontrados na imagem devem ser compactos, com todos os pixels agrupados em volta da média, ocupando uma região esférica no espaço. Um valor de desvio-padrão deve ser tomado como a medida desta compacidade (MATHER, 1987). Caso o desvio-padrão do agrupamento seja maior do que o valor estabelecido, o agrupamento que representa aquela classe de alvo é dividido. A outra condição é que os agrupamentos serão bem separados se seus centros são separados por uma distância mínima maior do que o valor de entrada. Se essa distância entre os centros de agrupamento for menor do que o valor de entrada, então os agrupamentos são fundidos. Divisão e fusão de agrupamento é um processo típico do ISODATA e, portanto, não mantém fixo uma quantidade de agrupamento. O processo de classificação começa com um número aleatório de médias de agrupamentos arbitrários uniformemente distribuídos no espaço de dados, os quais representam uma hipotética possibilidade de centros de agrupamentos. Então, iterativamente, as distâncias mínimas são calculadas para cada pixel da imagem e eles são reunidos nos agrupamentos mais próximos usando a técnica de distância mínima (menor distância ao centro dos agrupamentos). No fim dessa primeira sequência de alocação, a média dos valores de todos os pixels

alocados a cada agrupamento arbitrário é computada para cada uma das bandas, obtendo-se as coordenadas dos centros de cada *cluster*. Usando esses novos centros, uma nova iteração é realizada e a cada nova iteração é recalculada as médias, e os pixels são então reclassificados com relação às novas médias. As divisões ou fusões de clusters que irão ocorrer durante o processo de classificação dependem dos parâmetros dos padrões de entrada definidos pelo analista, que é o único momento em que ocorre uma participação interativa do analista com o processo de classificação. Alguns desses parâmetros, a exemplo do que são solicitados pelos softwares de processamento de imagens são:

- i) Número máximo de agrupamentos a serem considerados. Uma vez que cada agrupamento é a base de uma classe, esse número é o número máximo de classes a serem formadas. O método ISODATA começa determinando N médias de agrupamentos arbitrários. Alguns agrupamentos com poucos pixels podem ser eliminados;

- ii) Limiar (*threshold*) de convergência, que é a porcentagem máxima de pixels cujos valores de classe irão permanecer sem mudança entre iterações;

- iii) Um número mínimo de pixels em cada classe e caso houver um menor número de pixel na classe ela é desconsiderada e os pixels vão para a classe mais próxima;

- iv) Número de iterações; e;

- v) Se o desvio-padrão de uma classe for menor do que o limiar, então a classe é dividida em duas classes (WEILER, 2002). Em alguns casos alguns pixels podem ser não classificados, se eles não concordam com os parâmetros selecionados. A classificação finalizará quando o máximo de número de iterações for alcançado.

Essa classificação tem grande sucesso em descobrir os agrupamentos espectrais inerentes aos dados. Não importa onde estão os agrupamentos iniciais desde que um número suficiente de iterações seja permitido. As desvantagens do método são: o processo pode se repetir diversas vezes, consumindo tempo; o método não leva em consideração a homogeneidade espacial dos pixels; e se o número pré-definido de agrupamentos for alto, eles terão de ser fundidos, o que é um gasto desnecessário de processamento. Por isso, os algoritmos ISODATA limitam um número máximo de agrupamentos que o usuário deve definir inicialmente.

2.2.1.2 Método K-Means

A diferença da classificação K-means com o classificador ISODATA é que o primeiro assume que a quantidade de agrupamentos é conhecida *a priori*. Ele fixa um número de agrupamento, enquanto o ISODATA não mantém fixo este número. O objetivo do K-means é minimizar a variabilidade interna do agrupamento que é dimensionada pela função soma das distâncias ao quadrado (erros) entre cada pixel e o centro ao qual é alocado (PANDYA et al., 2013). A soma da distância ao quadrado (SS) tem a forma da equação 1.

$$SS_{dist} = \sum_{xy} [x - C(x)]^2 \quad (1)$$

onde $C(x)$ é a média do *cluster* ao qual o pixel x é alocado. É um critério para se avaliar a qualidade do *cluster*, que não se encontra no ISODATA. No restante, o K-means segue os mesmos procedimentos que foram apresentados para o ISODATA, mas como é mantida fixa uma quantidade arbitrária de agrupamentos, a divisão e a fusão de *clusters* não ocorrem.

2.2.2 Métodos de Classificação Supervisionados

O princípio da classificação supervisionada é baseado no uso de algoritmos para se determinar os pixels que representam valores de reflexão características para uma determinada classe. A classificação supervisionada é a mais utilizada na análise quantitativa dos dados de detecção remota.

Numa classificação supervisionada a primeira tarefa a ser realizada é a de selecionar conjuntos de pixels capazes de descrever espectralmente as classes que se quer classificar, dando-lhes o nome de áreas de treino.

Essa seleção poderá ser efetuada quer por fotointerpretação, quer por consulta de mapas ou através de visitas de campo. Estas amostras servem posteriormente para guiar os classificadores na atribuição dos pixels da imagem às respectivas classes.

Podemos resumir o processo de classificação supervisionada em três passos principais:

- Seleção de áreas de treino;
- Classificação da imagem;
- Avaliação de resultados.

A escolha das áreas de treino requer atenção redobrada, pois da qualidade espectral dessas amostras dependem os resultados dos classificadores. Existe uma relação delicada entre a importância da homogeneidade das áreas de treino de uma determinada classe e a sua variabilidade espectral, que garantirá a aproximação à natureza da classe. A seleção de áreas de treino é por si só um processo lento, implicando assim uma ação iterativa de escolha e análise da informação estatística das áreas, com o objetivo de alcançar a um patamar de separabilidade espectral, suficiente para garantir resultados satisfatórios na classificação.

Em seguida, o operador indica ao classificador as áreas de treino selecionadas, representativas das classes a classificar. Os resultados dependerão naturalmente da qualidade das áreas de teste e da natureza do classificador (FOODY, 2002).

Por último, e depois de terminada a classificação, o operador terá que avaliar a exatidão dos resultados para garantir a qualidade dos mesmos. Para isso, é necessário comparar quantitativamente os resultados obtidos com uma referência ou “verdade de campo” através da elaboração de um mapa classificado manualmente com as classes definidas. Os métodos de análise mais utilizados para a avaliação de resultados são a matriz de confusão, que compara as classes classificadas com o mapa de referência, e o índice de concordância Kappa (CONGALTON, 1991).

2.2.2.1 Classificação por Máxima Verossimilhança

A Máxima Verossimilhança (MAXVER) é o método tradicional de classificação comumente mais utilizado quando se necessita obter classes temáticas a partir de imagens de sensoriamento remoto. A distribuição espectral dessas classes é considerada Gaussiana ou normal, ou seja, objetos pertencentes à mesma classe irão apresentar resposta espectral próxima à média de valores daquela classe (MASELLI et al., 1992; RICHARDS, 1999). O MAXVER é um método de classificação supervisionado derivado do teorema de Bayes, que diz que a

distribuição *a posteriori* $P(i|\omega)$, isto é, a probabilidade para que um pixel com vetor de características ω pertença à classe i , é dada pela equação 2.

$$P(i|\omega) = \frac{P(\omega|i)P(i)}{P(\omega)} \quad (2)$$

onde $P(\omega|i)$ é função de probabilidade, $P(i)$ é a informação *a priori*, isto é, a probabilidade da classe i ocorrer na área de estudo e $P(\omega)$ é a probabilidade para que ω seja observado, e pode ser dada pela equação 3.

$$P(\omega) = \sum_{i=1}^M P(\omega|i)P(i) \quad (3)$$

onde M é o número de classes. $P(\omega)$ é geralmente tratado como uma constante de normalização para assegurar que o somatório $\sum_{i=1}^M P(i|\omega)$ seja igual a 1. O pixel x é atribuído a classe i pela regra dada pela equação 4.

$$x \in i \text{ se } P(i|\omega) > P(j|\omega) \text{ para todo } j \neq i \quad (4)$$

A MAXVER sempre assume que a distribuição de dados numa dada classe i obedece a uma distribuição Gaussiana multivariada. É então conveniente definir a probabilidade logarítmica *log* (ou função discriminante) pela equação 5.

$$g_i(\omega) = \ln P(\omega|i) = -\frac{1}{2}(\omega - \mu_i)' C_i^{-1}(\omega - \mu_i) - \frac{N}{2} \ln(2\pi) - \frac{1}{2} \ln(|C_i|) \quad (5)$$

Visto que *log* é uma função monótona¹, a equação (3) é equivalente a equação 6.

$$x \in i \text{ se } g_i(\omega) > g_j(\omega) \text{ para todos } j \neq i \quad (6)$$

com $-\frac{N}{2} \ln(2\pi) = \text{constante}$

¹ Uma função é monótona quando ela preserva ou inverte a relação de ordem entre dois conjuntos ordenados. Quando a função preserva a relação, ela é crescente, caso contrário, a função é decrescente (CHRISTOPHER e JAMES, 2014).

Cada pixel é atribuído à classe com a maior probabilidade ou etiquetado como não classificado se os valores da probabilidade estiverem todos abaixo de um limiar estabelecido pelo usuário (LILLESAND; KIEFER; CHIPMAN, 2004).

Os procedimentos gerais para MAXVER são os seguintes:

1. Determinação do número de tipos de cobertura de terra dentro da área de estudo.
2. Escolha das amostras de treinamento (pixels) para cada uma das classes desejadas usando a informação das coberturas de terra da área de estudo. Para o propósito, a distância Jeffries-Matusita (JM) pode ser usada para medir a separabilidade das classes das amostras de treinamento.
3. Uso das amostras treinadas para estimar o vetor média e matriz covariância de cada classe.
4. E, finalmente, é realizada a classificação de cada pixel na imagem em um dos tipos de cobertura de terra desejado ou etiquetado como desconhecido.

Este método baseia-se no princípio de que a classificação incorreta de um pixel não tem mais significado do que a classificação incorreta de qualquer outro pixel na imagem (RICHARDS, 1999). O método considera as distâncias médias ponderadas, utilizando parâmetros estatísticos para a distribuição dos pixels dentro de uma determinada classe (RICHARDS, 1986; CRÓSTA, 1993).

A aplicabilidade dos métodos de classificação tradicional como o MAXVER ou vizinho mais próximo pode ser limitada quando for usado um pequeno número de amostras de treinamento e espaço de características de alta dimensionalidade para dados hiperespectrais (LANDGREBE, 2003). Para obter um bom resultado com estes classificadores, é necessário escolher um número relativamente elevado de pixels para cada amostra de classe de treinamento, e que têm uma distribuição estatística mais próxima da distribuição normal (CRÓSTA, 1993; LANDGREBE, 2003).

Outro fator importante para sua efetividade é a obtenção de uma acurácia razoável da estimativa do vetor médio e da matriz de covariância de toda a classe espectral. Este fato é dependente da quantidade de pixels incluídos nas amostras de treinamento (RICHARDS, 1986; PEREIRA, 1996; SULSOFT, 2003). Este método de classificação assume que o usuário conhece de antemão a área analisada, juntamente com a distribuição de classes, de modo que quando a classificação é

realizada, a seleção das amostras de treinamento pode ser o mais eficiente possível (CRÓSTA, 1993).

2.2.2.1.1 Análise da classificação Máxima Verossimilhança

Segundo Richards (1999), para classes obedecendo a distribuição normal, a medida de separabilidade, neste caso, a distância de Jeffries - Matusita (JM) para duas classes, J_{ij} , é definida segundo equação 7.

$$J_{ij} = \sqrt{2(1 - e^{-\alpha})} \quad (7)$$

onde α é a distância de Bhattacharya, dada por:

$$\alpha = \frac{1}{8} (\mu_i - \mu_j)^t \left[\frac{(C_i + C_j)}{2} \right]^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left(\frac{\frac{|C_i + C_j|}{2}}{\sqrt{|C_i| |C_j|}} \right) \quad (8)$$

J_{ij} varia de 0 a 2, onde $J_{ij} > 1,9$ indica boa separabilidade das classes, separabilidade moderada para $1,0 \leq J_{ij} \leq 1,9$ e péssima separabilidade para $J_{ij} < 1,0$ (ENVI, 2006).

Na classificação MAXVER, cada classe é incluída numa região no espaço multispectral onde sua função discriminante é maior que a das outras classes. Essas regiões de classe são separadas por fronteiras de decisão, onde, a fronteira de decisão entre a classe i e j ocorre quando as funções se igualam, segundo equação 9.

$$g_i(\omega) = g_j(\omega) \quad (9)$$

Para distribuições normais multivariadas, a expressão em 9 se torna segundo é mostrado na equação 10.

$$-\frac{1}{2}(\omega - \mu_i)^t C_i^{-1}(\omega - \mu_i) - \frac{N}{2} \ln(2\pi) - \frac{1}{2} \ln(|C_i|) - \left(-\frac{1}{2}(\omega - \mu_j)^t C_j^{-1}(\omega - \mu_j) - \frac{N}{2} \ln(2\pi) - \frac{1}{2} \ln(|C_j|) \right) = 0 \quad (10)$$

Esta equação ainda pode ser reescrita como conforme equação 11.

$$-(\omega - \mu_i)^t C_i^{-1}(\omega - \mu_i) - \ln(|C_i|) + (\omega - \mu_j)^t C_j^{-1}(\omega - \mu_j) + \ln(|C_j|) = 0 \quad (11)$$

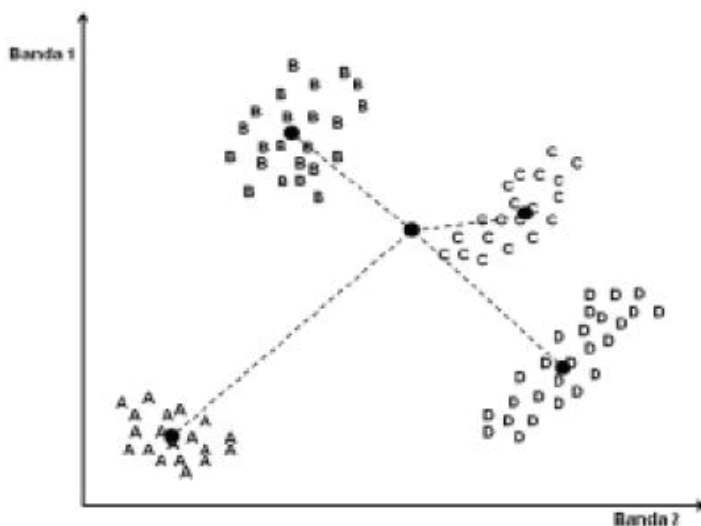
Essa é uma função quadrática em N dimensões. Contudo, se considerar apenas duas classes, as fronteiras de decisão são seções cônicas (isto é, parábolas, círculos, elipses ou hiperboles).

2.2.2.2 Método da Distância Mínima

O classificador de Distância Mínima é um algoritmo simples e rápido que não depende do número de pixels selecionados para cada área de treinamento, precisando somente de calcular as médias associadas a cada parcela de treino.

Cada área de treinamento é representada num gráfico de dispersão n -dimensional como uma nuvem de pixels (Figura 2). O algoritmo calcula para cada uma das áreas de treinamento, uma média que representa o centro geométrico desse conjunto de pixels a que se dá o nome de centroide. Posteriormente calcula a distância euclidiana entre todos os pixels da imagem e os centroides das áreas de treinamento, classificando o pixel com a classe que corresponder à menor distância calculada (CENTENO, 2009; CRISÓGONO, 2011).

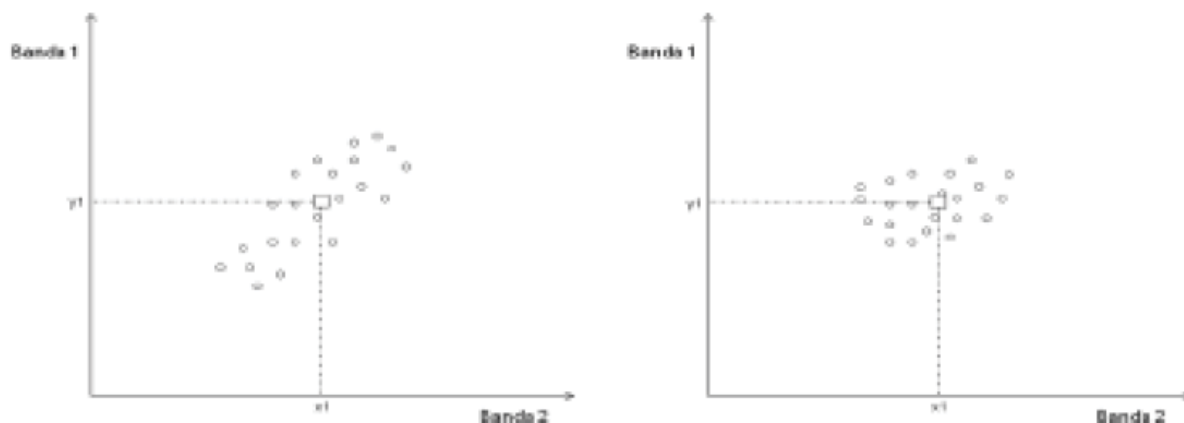
FIGURA 2 - ESQUEMA DO MÉTODO DE CLASSIFICAÇÃO DISTÂNCIA MÍNIMA.



FONTE: Adaptado de CRISÓGONO, 2011.

Apesar de muito rápido, o classificador baseia-se apenas na informação espectral das áreas de treinamento. Dado não ter em conta as covariâncias das áreas de treinamento, este classificador trata essas amostras, no seu domínio espectral, como simétricas, isto é, não tem em consideração, por exemplo, uma amostra espectralmente alongada (Figura 3).

FIGURA 3 - DUAS AMOSTRAS ESPECTRAIS COM MESMO CENTROIDE E DISPERSÕES DIFERENTES.



FONTE: Adaptado de CRISÓGONO, 2011.

Neste método, que considera a classe mais provável com base na distância espectral, é realizado o cálculo da média dos pixels em cada banda, para todas as classes e também a distância de um pixel à média de cada banda por classe. Logo, quanto maior a diferença espectral entre o pixel e a classe, maior a distância entre eles, portanto serão mais facilmente separados. Se o pixel for idêntico à média de uma classe, então a distância espectral entre o pixel e esta classe é nula (RIBEIRO et al., 2009; CENTENO, 2009). Uma desvantagem deste tipo de cálculo deve-se à dispersão dos valores de refletância em torno das médias.

O cálculo da Distância Mínima é efetuado através da equação 12.

$$D(x, w_j) = \sqrt{\sum_{i=1}^m (x_i - m_{ij})^2} \quad (12)$$

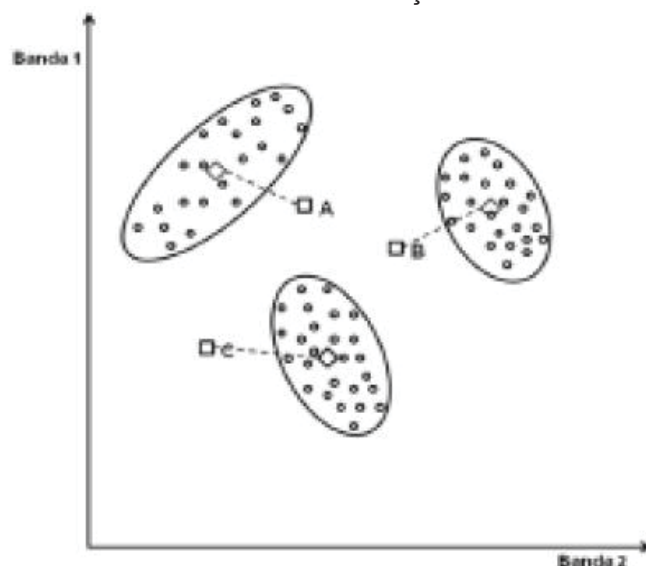
onde $D(x, w_j)$ é a diferença espectral; m_{ij} – é o valor da média da classe j na banda i .

2.2.2.3 Método da Distância de Mahalanobis

O classificador de Distância de Mahalanobis tem por base o classificador Distância Mínima, mas, ao contrário deste, não considera a existência de uma simetria na representação espectral das áreas de treinamento num gráfico de dispersão, isto é, não existe variabilidade nas áreas de treinamento nas n -dimensões da imagem (MENESES et al., 2012).

A Distância de Mahalanobis incorpora na Distância Euclidiana a dispersão das áreas de treinamento no espaço n -dimensional da imagem, ou seja, tem em conta a forma da nuvem de pixels das áreas de treinamento (Figura 4), na forma de matrizes de variância-covariância, conseguindo assim considerar a orientação e dispersão dos dados (RICHARDS; JIA, 1999; MENESES et al., 2012).

FIGURA 4 - ESQUEMA DO MÉTODO DE CLASSIFICAÇÃO DA DISTÂNCIA DE MAHALANOBIS.



FONTE: MENESES et al., 2012.

A Distância de Mahalanobis é, a medida da distância do pixel na posição x do espaço multidimensional ao centro da classe, dividida pela distância normalizada. Isso tem a propriedade de minimizar a distância do ponto ao centro de média. É dessa forma um classificador de distância sensitivo à direção, o que o faz similar ao classificador por Máxima Verossimilhança, mas assume que todas as classes tem covariâncias iguais. Isso lhe dá a vantagem de ser um classificador mais rápido do que o de Máxima Verossimilhança (RICHARDS; JIA, 1999; MENESES et al., 2012).

Para se usar a Distância de Mahalanobis para classificar um pixel para uma das n classes, inicialmente calcula-se a matriz de covariância com base nas amostras de treinamento das n classes, e o pixel será destinado à classe na qual a Distância de Mahalanobis seja a menor de todas. Usando a interpretação probabilística, comparativamente isso equivale à classe que apresenta a Máxima Verossimilhança (RICHARDS; JIA, 1999; MENESES et al., 2012).

A medida da Distância de Mahalanobis considera que um pixel tem uma distância menor ao centro de média que um outro pixel candidato, se ele se distribuir ao longo do eixo da classe com maior variabilidade. Ele deve, então, ser classificado

a essa classe. As medidas Mahalanobis computadas definem um espaço multidimensional cujos limites determinam o intervalo de variação tido por aceitável para que pixels candidatos possam ser classificados como pertencentes a uma das classes existentes (MENESES et al., 2012).

A medida de Mahalanobis é usada para identificar pixels que se distanciam bastante dos pixels da distribuição. Muitas vezes, quando são selecionadas as áreas de amostragens das classes de treinamento, alguns pixels podem contaminar a homogeneidade das classes, levando a erros nos cálculos da média e da matriz de covariância. Técnicas de regressão podem ser utilizadas para determinar se um pixel é atípico ou não no espaço multivariado, usando medidas da distância quadrática, que classifica os pixels na base da menor das distâncias $d(x, m_i)$ sensível à direção, na forma de (RICHARDS; JIA, 1999; CENTENO, 2009):

$$d(x, m_i)^2 = (x - m_i)^t C^{-1} (x - m_i) \quad (13)$$

2.3 MÉTODOS DE REDUÇÃO DE DIMENSIONALIDADE DE IMAGENS HIPERESPECTRAIS

Dados de sensoriamento remoto muitas vezes apresentam alta dimensionalidade. A fim de lidar com tais dados do mundo real de forma adequada, sua dimensionalidade precisa ser reduzida. Redução de dimensionalidade é a transformação de dados de alta dimensão em uma representação significativa de dimensionalidade reduzida. Idealmente, a representação reduzida deveria ter uma dimensionalidade que corresponda à dimensionalidade intrínseca dos dados. A dimensionalidade intrínseca dos dados é o número mínimo de parâmetros necessários para levar em consideração as propriedades observadas dos dados (FUKUNAGA, 1990). A redução de dimensionalidade é importante em muitos domínios, uma vez que mitiga o problema da dimensionalidade e outras propriedades indesejáveis em espaços de alta dimensão (JIMENEZ; LANDGREBE, 1997). Como resultado, a redução de dimensionalidade facilita, entre outros, a classificação, a visualização e a compressão de dados de alta dimensão.

2.3.1 Propagação de Afinidade

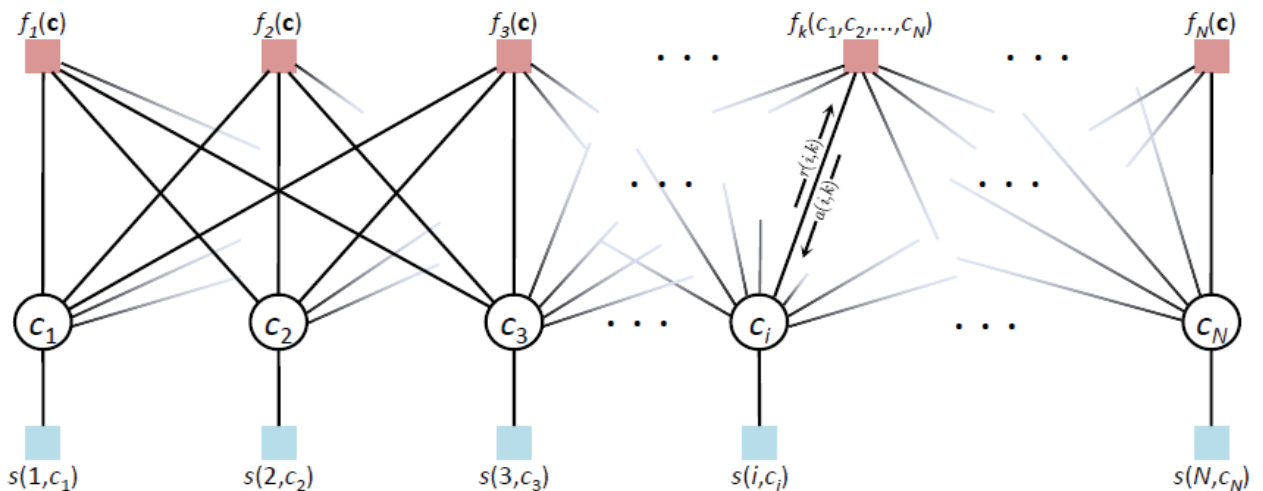
A Propagação de Afinidade (PA) é um algoritmo de *clusterização* criado por FREY e DUECK em 2007. Este algoritmo identifica centros de *clusters*, também referidos como exemplares, e forma clusters de pontos em torno desses exemplares. A PA opera considerando simultaneamente todos os pontos como potenciais exemplares, e efetua troca de mensagens entre os pontos até que surja uma boa disposição de exemplares e *clusters* (FREY; DUECK, 2007).

Como dado de entrada, a PA usa o valor real de similaridades, $S(i, j)$, que descrevem como o j -ésimo ponto é adequado para servir como exemplar para o i -ésimo ponto. Para $i = j$, a similaridade de um ponto i , $S(i, i)$ indica quão provável o i -ésimo ponto é para ser escolhido como um exemplar, e é chamado de preferência. Normalmente, a similaridade é uma distância Euclidiana Quadrática Negativa definida conforme a equação 14. Os parâmetros x_i e x_j são valores digitais i e j dos vetores de estado do elemento (*clusterização* supervisionada) ou bandas i e j da imagem, com dimensão espacial M, N no espaço espectral.

$$S(i, j) = -\|x_i - x_j\|^2 = -\sum_{m, n=1}^{M, N} (x_{mni} - x_{mnj})^2 \quad (14)$$

O número de exemplares selecionados, isto é, os *clusters* definidos, é influenciado pelos valores de preferência, mas também surge do processo de envio de mensagem no gráfico de fatores mostrado na Figura 5.

FIGURA 5 - GRÁFICO DE FATORES DA PROPAGAÇÃO DE AFINIDADE.

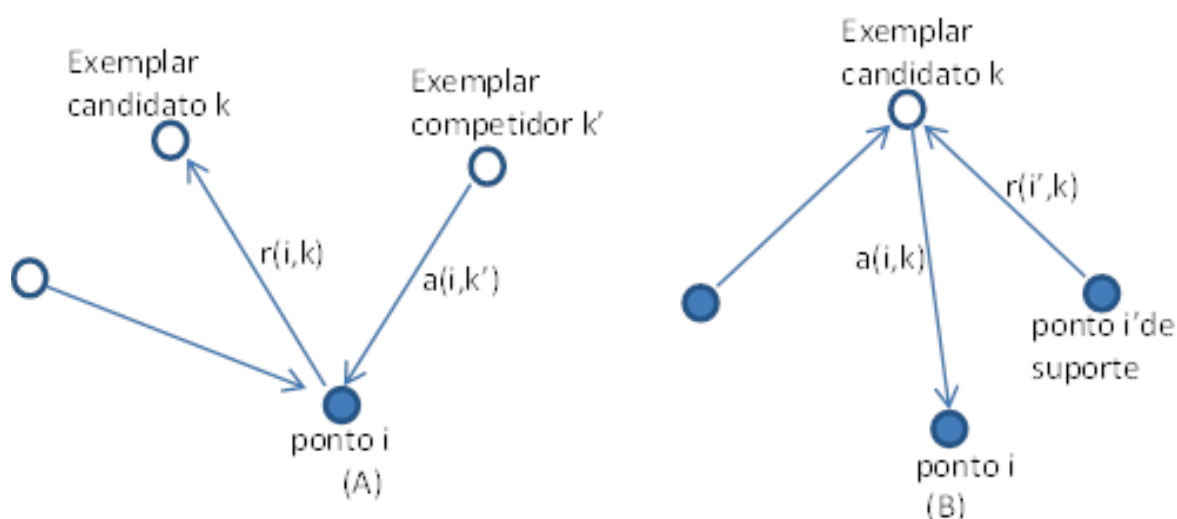


FONTE: FREY; DUECK, 2007

De acordo com a figura, cada componente da função global é representado por um nó de função e cada c_i é representado por um nó de variável. Cada termo $f_k(c)$ no gráfico tem um nó de função correspondente que está conectado a todas as variáveis c_1, c_2, \dots, c_N . Além disso, cada termo $s(i, c_i)$ tem um nó de função correspondente, conectado a única variável c_i . O \log da função global, nesse caso a similaridade $S(i, j)$ da rede é dado pela soma de todas funções- \log representadas por nós de função.

O processo de envio de mensagens no gráfico de fatores é baseado em dois tipos de mensagens, responsabilidades e disponibilidades. A responsabilidade enviada a partir do ponto i ao exemplar candidato ponto k , reflete a evidência acumulada de quão bem adequado o ponto k é para servir como exemplar para o ponto i , tendo em conta outros potenciais exemplares para o ponto i . A disponibilidade, enviada a partir do exemplar candidato ponto k para o ponto i , reflete a evidência acumulada de quão adequado seria para o ponto i escolher o ponto k como seu exemplar, tendo em conta o apoio de outros pontos de que ponto k deve ser um exemplar (DUECK, 2009). O processo de troca de mensagens de responsabilidade e de disponibilidade é mostrado na Figura 6.

FIGURA 6 - DUAS MENSAGENS SÃO TROCADAS ENTRE PONTOS: (A) “RESPONSABILIDADES” $R(i, k)$ SÃO ENVIADAS DO PONTO i PARA O EXEMPLAR CANDIDATO k , E (B) “DISPONIBILIDADES” $A(i, k)$ SÃO ENVIADAS DO EXEMPLAR CANDIDATO k PARA O PONTO i .



FONTE: DUECK, 2009.

Durante o processo de propagação de mensagens, acontecem oscilações numéricas devido ao grande fluxo de mensagens. Contudo, ao atualizar as mensagens, é importante que estas sejam amortecidas para evitar oscilações que

possam surgir em determinadas circunstâncias. Para tal, é usado o fator de amortecimento (*damping factor*) λ , que varia entre 0 e 1. Este fator λ , deve ser igual a pelo menos 0,5 e menor que 1. Determinado o valor de λ , cada mensagem amortecida é definida ao multiplicar λ pelo valor da referida mensagem na iteração anterior, adicionado do produto entre $(1 - \lambda)$ e o valor atualizado da mensagem (FREY; DUECK, 2007).

Cada iteração da Propagação de Afinidade consiste na atualização de todas as responsabilidades dadas as disponibilidades, atualização de todas as disponibilidades dadas as responsabilidades, e na combinação das disponibilidades e responsabilidades para monitorar as decisões dos exemplares e terminar o processo de *clusterização*.

Os valores da responsabilidade e da disponibilidade são ajustados da seguinte forma:

$$\forall_{i,k} : r(i,k) = S(i,k) - \max_{k' : k' \neq k} [S(i,k') + a(i,k')] \quad (15)$$

$$\begin{aligned} \forall_{i,k} : a(i,k) &= \sum \max[0, r(i',k)], \quad \text{para } k = i, \quad i' : i' \neq i \\ \forall_{i,k} : a(i,k) &= \min \left[0, r(k,k) + \sum_{i' : i' \in \{i,k\}} \max[0, r(i',k)] \right], \quad \text{para } k \neq i \end{aligned} \quad (16)$$

Na equação 15, a letra i representa um ponto e k' representa um exemplar candidato competidor. Na primeira iteração, as disponibilidades são inicializadas em zero, e $r(i,k)$ é disposto para a similaridade de entrada entre o ponto i e ponto k como seu exemplar, subtraído o máximo das similaridades entre o ponto i e outros exemplares candidatos k' .

Nas equações 16, a disponibilidade $a(i,k)$ é estabelecida para responsabilidade própria (*self-responsibility*) $r(k,k)$ adicionada a soma das responsabilidades positivas que o exemplar candidato k recebe de outros pontos i' que o apoiam.

A cada iteração, a atribuição de pontos aos exemplares é definida como:

$$\phi(x_i) = \arg \max_k \{r(i,k) + a(i,k)\} \quad (17)$$

Na equação acima, $\phi(x_i)$ denota o exemplar para o ponto ou banda x_i . Em cada ponto, a soma das matrizes de responsabilidade $r(i,k)$ e de disponibilidade $a(i,k)$ fornece a informação necessária da *clusterização* do ponto i . O k com o máximo da soma $r(i, k) + a(i, k)$ representa exemplar do ponto i .

O processo de troca de mensagens para após um determinado número de iterações ou assim que a estrutura do *cluster* não altere significativamente para um dado número de iterações, isto é, o processo converge se todos os exemplares $\phi(x_i)$ permanecem inalterados para algumas iterações contantes, geralmente 10 iterações (Dueck, 2009).

2.3.2 Análise de Componentes Principais

A Análise de Componentes Principais (PCA) (do inglês *Principal Component Analysis*) foi inventada em 1901 por KARL PEARSON em analogia ao teorema dos eixos principais em mecânica (PEARSON, 1901). Mais tarde foi desenvolvida independentemente e nomeada por Harold Hotelling em 1930 (HOTELLING, 1933).

A PCA é um procedimento estatístico multivariado que transforma um conjunto de observações de variáveis correlacionadas em outro conjunto de valores de variáveis linearmente não correlacionadas, designadas componentes principais.

As componentes principais apresentam propriedades importantes: cada componente principal é uma combinação linear de todas as variáveis originais, independentes entre si e estimadas com o propósito de reter, em ordem de estimação, o máximo de informação, em termos da variabilidade total contida nos dados.

A PCA está associada à idéia de redução de dimensionalidade de dados, com menor perda possível da informação. Procura-se redistribuir a variação observada nos eixos originais de forma a se obter um conjunto de eixos ortogonais não correlacionados.

O número de componentes principais distintos é igual ao menor número das variáveis originais ou ao número de observações menos um. Esta transformação é definida de tal forma que a primeira componente principal tenha a maior variância possível (isto é, representa a variabilidade dos dados o quanto possível), e cada uma das componentes sucessivas, por sua vez, tem a maior variância possível sob o constrangimento de ser ortogonal a componentes anteriores. Esta técnica pode ser

utilizada para geração de índices e agrupamento de indivíduos (SOLOMON; BRECKON, 2011).

A análise agrupa os indivíduos de acordo com sua variação, isto é, os indivíduos são agrupados segundo suas variâncias, ou seja, segundo seu comportamento dentro da população, representado pela variação do conjunto de características que define o indivíduo, ou seja, a técnica agrupa os indivíduos de uma população segundo a variação de suas características.

Os passos para calcular as componentes principais são:

- 1) Obter os dados ou as M amostras de vetores de dimensão n ;
- 2) Calcular a média ou o vetor médio destes dados;
- 3) Subtrair a média de todos os itens de dados;
- 4) Calcular a matriz de covariância usando todas as subtrações. Ela é o resultado da média do produto de cada subtração por ela mesma e terá dimensão $n \times n$;
- 5) Calcular os autovalores e autovetores da matriz de covariância.
- 6) Arranjar a matriz de Transformada de Hotelling², cujas linhas são formadas a partir dos autovetores da matriz de covariância arranjados de modo que a primeira linha, o elemento $(0,0)$, seja o autovetor correspondente ao maior autovalor, e assim sucessivamente até que a última linha corresponda ao menor autovalor.

A transformação $T = X W$ mapeia um vetor de dados $x_{(i)}$ a partir de um espaço original de p variáveis para um novo espaço de p variáveis não correlacionadas sobre o conjunto de dados. No entanto, nem todas as componentes principais precisam ser mantidas. Manter somente os primeiros componentes principais L , produzidos usando somente os primeiros L vetores de carregamento, dá a seguinte transformação:

$$T_L = X W_L \quad (18)$$

² Transformada de Hotelling é uma representação de um processo estocástico como uma combinação linear infinita de funções ortogonais. É análogo a uma representação de séries de Fourier de uma função em um intervalo limitado. Esta transformação está intimamente relacionada com a análise de componentes principais (SACHIN, 2011).

onde, a matriz T_L agora tem n linhas, mas apenas L colunas. Em outras palavras, PCA obtém uma transformação linear:

$$t = W^T x, x \in R^p, t \in R^L \quad (19)$$

onde as colunas $p \times L$ da matriz W formam uma base ortogonal para as L características (as componentes de representação t) que são decorrelacionadas (BENGIO; COURVILLE; VINCENT, 2013). Pela construção de todas as matrizes de dados transformadas com apenas L colunas, esta matriz de pontuação maximiza a variância nos dados originais que foram preservados, enquanto minimiza o erro total quadrático de reconstrução, dado por:

$$\|TW^T - T_L W_L^T\|_2^2 \quad \text{ou} \quad \|X - X_L\|_2^2 \quad (20)$$

Tal redução de dimensionalidade pode ser um passo muito útil para visualização e processamento de dados de alta dimensionalidade, enquanto ainda mantem o máximo da variância nos dados o quanto possível. No caso em que as duas direções através dos dados (ou duas das variáveis originais) são escolhidas de forma aleatória, os clusters podem ser muito menos dispersos uns dos outros, e pode ser de fato muito mais provável de se sobreporem substancialmente entre si, tornando-se indistinguíveis.

De forma semelhante, na análise de regressão, quanto maior o número de variáveis explicativas, maior é a chance de ajustar o modelo, produzindo conclusões que não generalizam outros conjuntos de dados. Uma abordagem, especialmente quando existem fortes correlações entre as diferentes possíveis variáveis explicativas, é reduzi-las a algumas componentes principais e, em seguida, executar a regressão neles, um método chamado regressão de componentes principais.

A redução de dimensionalidade também pode ser apropriada quando as variáveis no conjunto de dados são ruidosas. Se cada coluna do conjunto de dados contém ruído gaussiano independente igualmente distribuído, então as colunas de T também irão conter ruído gaussiano similar igualmente distribuído (tal distribuição é invariante sob os efeitos da matriz W , o que pode ser pensado como uma rotação de alta dimensão dos eixos coordenados). No entanto, com maior parte da variância

total concentrada nas primeiras componentes principais em comparação com a mesma variância de ruído, o efeito proporcional do ruído é menor, as poucas primeiras componentes alcançam uma alta taxa sinal-ruído. Assim, a PCA pode ter o efeito de concentrar muito do sinal para as poucas primeiras componentes principais, que podem ser utilmente capturadas por redução de dimensionalidade; enquanto que as últimas componentes principais podem ser dominadas pelo ruído, e assim eliminadas sem grandes perdas.

2.4 FILTRAGENS DE IMAGENS

Ruídos podem surgir na captura e no processo de transmissão da imagem. O ruído é normalmente dividido em ruído gaussiano, o ruído balanceado e o ruído impulsivo. Os ruídos de impulso surgem como luz e com os pixels de ruído escuros distribuídos aleatoriamente sobre a imagem. Isto não só corrompe a verdadeira informação da imagem, mas também afeta seriamente os efeitos visuais da imagem. Portanto, a redução de ruídos de impulso tem importante significado para a área de processamento e análise de imagens e na visão computacional (GONZALES; WOODS, 2000).

Para uma imagem corrompida por ruídos, podemos usar métodos de filtragem linear ou não linear para redução de ruídos. No domínio da frequência, os detalhes são componentes de alta frequência da imagem, que podem ser facilmente confundidos com ruídos de alta frequência. Portanto, como manter os detalhes da imagem e filtrar efectivamente ruídos aleatórios é a chave para o processo de filtragem de imagens (ZHU; HUANG, 2012). Para o efeito são usados filtros de passa-baixa que preservam os componentes de baixa frequência da imagem a custo de reduzir os de alta frequência. Dentre as técnicas mais conhecidas de suavização estão os filtros da média e da mediana (MENESES et al., 2012).

2.4.1 Filtragem pela Mediana

O filtro da mediana é uma metodologia de processamento de sinais não linear com base em estatísticas. É uma alternativa para minimizar um pouco a aparência borrada da imagem pelo filtro da média, e muitas vezes usado para remover o ruído de uma imagem ou sinal (MENESES et al., 2012). Essa redução de ruído é uma

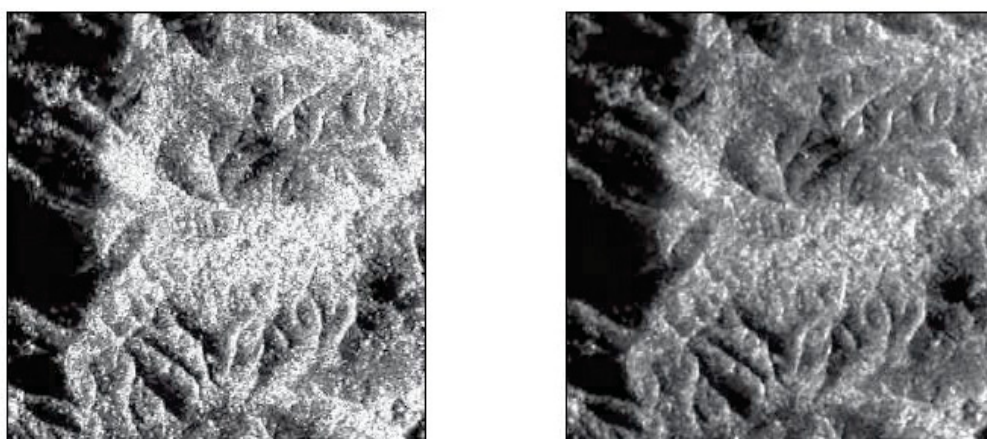
etapa típica de pré-processamento para melhorar os resultados do processamento posterior, como por exemplo, a detecção de bordas em uma imagem (GONZALES; WOODS, 2000). A filtragem mediana é amplamente utilizada em processamento digital de imagens porque, sob certas condições, preserva bordas enquanto remove o ruído. O valor ruidoso da imagem digital ou a sequência é substituído pelo valor da mediana da vizinhança (máscara). Os pixels da máscara são classificados na ordem dos seus níveis de cinza, e o valor da mediana do grupo é guardado para substituir os valores ruidosos (ZHU; HUANG, 2012). A saída do filtro da mediana é dada pela equação 21.

$$g(x,y) = med \{f(x-i,y-i), \quad i,j \in W\} \quad (21)$$

onde, $f(x,y)$ e $g(x,y)$ são a imagem original e a imagem de saída respectivamente, W é a máscara bidimensional. O tamanho da máscara é $n \times n$ (onde n é geralmente ímpar) tal como 3×3 , 5×5 , e etc.; a forma da máscara pode ser quadrada, circular, cruz, e etc (ZHU; HUANG, 2012).

O filtro da mediana é um filtro passa-baixa, preservando os componentes de baixa frequência da imagem à custa da redução dos de alta frequência. A Figura 7 mostra os efeitos do filtro da mediana sobre uma imagem.

FIGURA 7 - EXEMPLO DO EFEITO DE ATENUAÇÃO DE RUÍDO EM UMA IMAGEM DE RADAR, (A) ANTES E (B) DEPOIS DO USO DO FILTRO DA MEDIANA.



(a) FONTE: MENESES et al., 2012.

(b)

2.5 AVALIAÇÃO DA ACURÁCIA E VALIDAÇÃO DA CLASSIFICAÇÃO

A avaliação da acurácia é uma parte importante em qualquer projeto de classificação. No contexto de extração de informações através de análise de imagem, a acurácia "mede a concordância entre um padrão assumido ser correto e uma imagem classificada de qualidade desconhecida" (CAMPBELL, 2007). A precisão define o nível de detalhe encontrado dentro da classificação. É possível aumentar a acurácia de uma classificação pela diminuição da quantidade de detalhe ou pela generalização de classes amplas, em vez de muito específicas. Por exemplo, um esquema que geralmente categoriza árvores *versus* culturas apresenta menor possibilidade de erro na classificação do que um que distingue vários tipos de árvores e vários tipos de culturas. Neste caso, menor precisão oferece o potencial para maior acurácia. No entanto, o usuário do mapa que oferece apenas classes gerais não pode fazer afirmações precisas sobre qualquer ponto do mapa (FOODY, 2002).

O erro de classificação ocorre quando um pixel (ou característica) pertencente a uma categoria é atribuído a outra categoria. Os erros de omissão ocorrem quando uma característica é deixada de fora da categoria sendo avaliada; erros de inclusão ocorrem quando uma característica é incorretamente incluída na categoria a ser avaliada. Um erro de omissão em uma categoria será contado como um erro de inclusão em outra categoria (CAMPBELL, 2002; FOODY, 2002).

Avaliação da acurácia é realizada comparando o mapa criado por análise de sensoriamento remoto a um mapa de referência com base em outra fonte de informação (LILLESAND; KIEFER; CHIPMAN, 2004). Podem ser feitas perguntas tais como por que é necessária a análise de sensoriamento remoto se o mapa de referência para comparar já existe. Um dos propósitos principais da avaliação da acurácia e análise de erros neste caso são de permitir comparações quantitativas de diferentes interpretações. Classificações feitas a partir de imagens adquiridas em momentos diferentes, classificadas por diferentes procedimentos, ou produzidas por indivíduos diferentes podem ser avaliadas usando uma comparação pixel por pixel, ponto-a-ponto. Os resultados devem ser considerados no contexto da aplicação para determinar qual é a "mais correta" ou "mais útil" para um propósito particular (FOODY, 2002).

A fim de ser comparado, tanto o mapa a ser avaliado e o mapa de referência devem ser geometricamente registrados um ao outro com acurácia. Eles também devem usar o mesmo esquema de classificação, e deveriam ter sido classificados no mesmo nível de detalhe. Um método simples de comparação é calcular a área total atribuída a cada categoria em ambos os mapas e comparar os valores globais. Esse tipo de avaliação é chamado acurácia específica não local. Por outro lado, a acurácia específica local é baseada na comparação dos dois mapas em locais específicos (ou seja, pixels individuais em duas imagens digitais) (CAMPBELL, 2002; FOODY, 2002; LILLESAND; KIEFER; CHIPMAN, 2004). Neste tipo de comparação, é óbvio que o grau no qual os pixels em uma imagem se alinham espacialmente com os pixels na segunda imagem contribui para o resultado da avaliação da acurácia. Erros na classificação devem ser distinguidos de erros no registo ou posicionamento de fronteiras. Outra forma útil de avaliação da acurácia específica local é comparar os dados de campo ou dados de treinamento em diversos locais dentro da imagem, semelhante à maneira como a avaliação da acurácia espacial é realizada utilizando pontos de verificação de campo para ortofotos e modelos digitais de terreno (LILLESAND; KIEFER; CHIPMAN, 2004).

A acurácia de classificação de imagem é mais frequentemente relatada como uma percentagem correta. A acurácia do usuário (AU) é calculada usando o número de pixels classificados corretamente e o número total de pixels atribuídos a uma determinada categoria (STORY; CONGALTON, 1986). Leva em conta os erros de inclusão, e é calculada dividindo cada um dos elementos da diagonal na matriz de confusão pelo total da linha em que tal ocorre, segundo a equação 22.

$$AU = \frac{x_{ii}}{x_{i\bullet}} \times 100\% \quad (22)$$

onde: x_{ii} – elemento na posição i-ésima linha e i-ésima coluna; $x_{i\bullet}$ - somatório da linha i .

A acurácia do produtor (APr) informa ao analista de imagens sobre o número de pixels corretamente classificados em uma categoria específica, como percentagem do número total de pixels realmente pertencentes a essa categoria na imagem. Acurácia do produtor mede os erros de omissão, e é dada pela expressão 23.

$$APr = \frac{x_{jj}}{x_{\bullet j}} \times 100\% \quad (23)$$

onde, $x_{\bullet j}$ - somatório da coluna j .

A matriz de confusão é representada pela acurácia global (AG), a acurácia do usuário (AU), a acurácia do produtor (APr) e pelo coeficiente kappa (CK). A AG representa a razão entre o número de amostras corretamente reconhecidas pelo algoritmo de classificação e o número total de amostras de teste, e é dada pela expressão 24.

$$AG = \frac{\sum_{j=1}^C x_{jj}}{N} \times 100\% \quad (24)$$

onde, C e N são respectivamente o número total de classes e de pixels. De acordo com SCEPAN (1999), a AG mínima aceitável é de 85%.

O coeficiente Kappa (CK) é uma medida que expressa a diferença entre a concordância observada entre dois mapas e o acordo que pode ser alcançado apenas por acaso, fazendo a correspondência dos dois mapas (CAMPBELL, 2002). Esta é a segunda medida da acurácia da classificação que incorpora os elementos fora da diagonal assim como os da diagonal da matriz de confusão, para dar uma avaliação mais robusta da acurácia que a acurácia global, e é calculado segundo equação 25 (JENSEN, 1996).

$$CK = \frac{\sum_{j=1}^C \frac{x_{jj}}{N} - \sum_{j=1}^C \frac{x_{j\bullet} \cdot x_{\bullet j}}{N^2}}{1 - \sum_{j=1}^C \frac{x_{j\bullet} \cdot x_{\bullet j}}{N^2}} \quad (25)$$

onde $x_{j\bullet}$ - somatório da linha j .

O valor do CK se encontra no intervalo $[-1, +1]$. Quanto mais o valor de CK se aproxima de $+1$, melhor é a classificação. Embora valores de CK negativos sejam possíveis, Cohen (1960) nota que são improváveis de acontecer na prática, e quando acontece é um indicador de um problema sério, porque valores negativos de CK representam uma discordância. Segundo Cohen (1960), Landis e Koch (1977), o coeficiente CK pode ser interpretado de acordo com a Tabela 1:

TABELA 1 - INTERPRETAÇÃO DOS VALORES DO COEFICIENTE KAPPA (CK).

Valores dos níveis de concordância do Coeficiente Kappa		
0-0,2	Nenhum	0-4%
0,21-0,39	Mínimo	4-15%
0,40-0,59	Fraco	15-35%
0,60-0,79	Moderado	35-63%
0,80-0,90	Forte	64-81%
>0,90	Quase perfeito	82-100%

FONTE: COHEN (1960), LANDIS e KOCH (1977).

Depois de observar os resultados da avaliação da acurácia, pode ser preciso ajustar as amostras de treinamento ou parâmetros de classificação, ou escolher um classificador diferente para obter um resultado melhor e validar a classificação. Se este for o caso, deve-se usar os dados da nova classificação como entrada para atualizar os pontos da avaliação da acurácia e calcular a matriz de confusão.

CAPÍTULO 3. MATERIAL E MÉTODOS

3.1 MATERIAL

O método proposto nesta tese foi implementado usando os seguintes materiais: Um computador HP Pavilion dv6 com processador Intel (R) Core (TM) i7 CPU 2,00 GHz e 8 GB de memória RAM; Software MATLAB R2016a para criação de rotinas para Propagação de Afinidade; ENVI 4.6.1 foi aplicado na remoção de bandas com ruído e informação redundante, filtragem e classificação das imagens usando a Máxima Verossimilhança e para avaliação da acurácia pelo uso da matriz de confusão e coeficiente Kappa; MultiSpec 3.4.1 foi usado nesta tese na coleta de amostras para *clusterização* supervisionada com a Propagação de Afinidade. O software ERDAS IMAGINE 2014 foi usado para apoiar na coleta de amostras de treinamento e de teste nas imagens hiperespectrais AVIRIS e HYDICE e para salvar as imagens criadas no ENVI em outros formatos possíveis de leitura no MATLAB. Os cálculos e as tabelas foram todos obtidos com base no Microsoft Excel 2010.

3.2 LOCALIZAÇÃO DA ÁREA DE ESTUDO E IMAGENS UTILIZADAS

Para avaliação e validação do método proposto nesta tese foram usadas duas disposições de dados hiperespectrais (AVIRIS e HYDICE), frequentemente usadas em pesquisas nas áreas de sensoriamento remoto e visão computacional (JIA; QIAN, 2009; QIAN et al., 2011). Essas imagens foram escolhidas por cobrirem dois cenários completamente diferentes, sendo o primeiro uma área agrícola rural e o outro uma área urbana, conforme ilustram as Figuras 8 e 9. Essa diferenciação de cenários, permite testar e validar a aplicabilidade e eficiência do algoritmo em cenas mais diversificadas e com diferentes níveis de complexidade.

As amostras para treinamento e teste do algoritmo, apresentadas nas Tabelas 2 e 3, foram coletadas das duas imagens, obedecendo uma distribuição quase uniforme e de modo que seus totais não ultrapassassem 5% do tamanho da imagem, o que segundo Arya, Antonisamy e Kumar (2012), evita possíveis tendenciosidades no processo de classificação.

3.2.1 Imagem AVIRIS

Esta é uma imagem hyperspectral 92AV3C, disponível em <http://www.tec.army.mil/Hypercube>, e foi coletada pelo sensor Airborne Visible Infrared Imaging Spectrometer (AVIRIS) em 1992, e corresponde ao local de ensaio de Indian Pine na região noroeste do estado de Indiana. Das 220 bandas que compõem a imagem, 20 (104-108, 150-163, 220) são ignoradas por causa do ruído (aquelas que cobrem a região de absorção de água ou com baixa taxa entre o sinal e o ruído) (BAUMGARDNER; BIEHL; LANDGREBE, 2015), e 10 foram excluídas por conter informação redundante, ficando a imagem final preprocessada com o total de 190 bandas espectrais. A imagem tem dimensão espacial de 145 x 145 pixels, correspondentes às linhas e colunas. A resolução espacial é de 20m por pixel.

A quantidade de pixels nas classes varia de 20 a 2468 pixels. Nelas, três diferentes estados de crescimento de soja podem ser encontrados, junto com outros três diferentes estágios de crescimento do milho. Bosques, grama/pasto e árvores são as maiores classes em termos de número de amostras (pixels). Classes pequenas incluem torres, feno-arado, alfafa, edifício-pasto-passadeiras, aveia, grama-pasto e trigo. No total, o conjunto de dados tem 16 classes rotuladas (LANDGREBE, 2003). A imagem hiperespectral AVIRIS e a verdade de campo usada para executar os experimentos são mostradas na Figura 8.

FIGURA 8 - IMAGEM AVIRIS DO LOCAL DE ESTUDO E VERDADE DE CAMPO.



FONTE: <http://www.tec.army.mil/Hypercube>.

TABELA 2 - CLASSES E VERDADES DE CAMPO PARA TREINAMENTO E TESTE PARA AVIRIS.

Nr.	Classe	Treinamento	Teste	Total
1	Alfafa	22	17	39
2	Milho em plantio direto	94	100	194
3	Milho para sementeira	60	108	168
4	Milho	48	16	64
5	Gramma/Pasto	68	50	118
6	Gramma/Árvores	82	50	132
7	Gramma/Pasto-cortado	20	9	29
8	Feno-arado	52	62	114
9	Aveia	16	5	21
10	Soja em plantio direto	80	96	176
11	Soja para sementeira	162	137	299
12	Soja limpa	54	78	132
13	Trigo	57	47	104
14	Bosque	102	160	262
15	Edificação-Grama-Árvore-Passadeiras	64	103	167
16	Torres de Pedra e Aço	47	30	77

FONTE: O AUTOR, 2018.

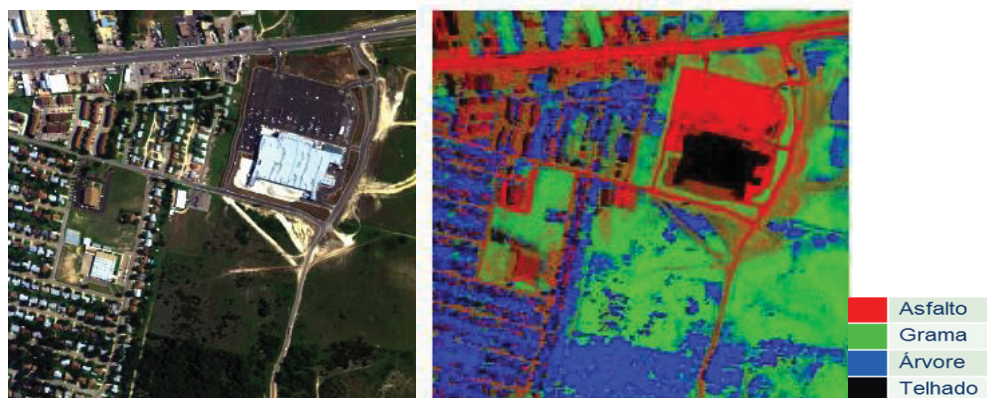
A imagem AVIRIS foi escolhida para essa pesquisa, apesar de apresentar uma área relativamente pequena, por se tratar de uma das melhores imagens hiperespectrais encontradas no domínio público, apresentando um número maior de classes e de bandas, características importantes para realização do presente estudo.

3.2.2 Imagem HYDICE

Dado HYDICE, disponível em <http://www.tec.army.mil/Hypercube>, é uma das imagens hiperespectrais mais utilizada em estudos de separação de mistura hiperespectral (LIU et al., 2011; QIAN et al., 2011). Esta imagem foi coletada em Outubro de 1995 pelo sensor HYDICE (Hyperspectral Digital Imagery Collection Experiment), e representa a localização de uma área urbana em Copperas Cove, nos Estados Unidos. Existem 307 x 307 pixels, cada um dos quais corresponde a uma área de 2 x 2 m². Nesta imagem, existem 210 bandas espectrais variando entre os comprimentos de onda de 400 nm a 2500 nm, o que resulta em uma largura espectral de banda de 10 nm. Foram removidas as bandas 1-4, 76, 87, 101-111, 136-153 e 198-210, devido ao vapor de água denso e efeitos atmosféricos, tendo a imagem preprocessada permanecido com uma resolução espectral de 162 bandas. Existe também uma verdade de campo com 4 classes, nomeadamente asfalto, grama, árvore e telhado. O asfalto inclui a estrada e a área de estacionamento; a

grama cobre áreas com aparência verde; árvore apresenta uma assinatura espectral diferente da grama, e aparece verde escuro principalmente nas áreas na zona inferior da imagem; telhado aparece branco na zona central e a esquerda da imagem, e alguns aparecem em marrom. A imagem original e a verdade de campo são mostrados na Figura 9.

FIGURA 9 - IMAGEM HYDICE DA ÁREA DE ESTUDO E VERDADE DE CAMPO.



FONTE: <http://www.tec.army.mil/Hypercube>.

TABELA 3 - CLASSES E VERDADES DE CAMPO PARA TREINAMENTO E TESTE PARA HYDICE.

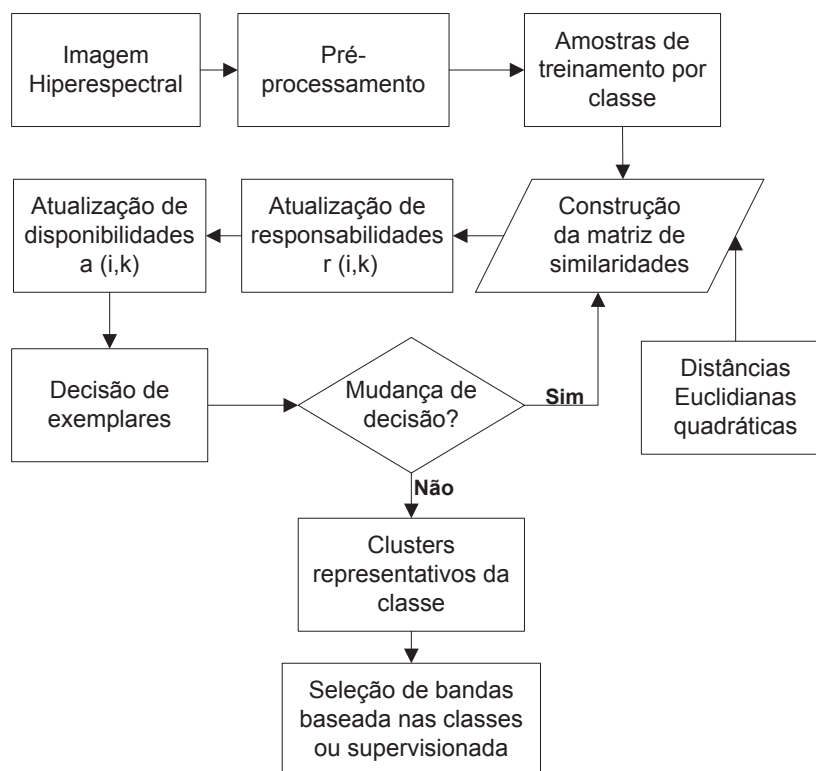
Nr.	Classe	Treinamento	Teste	Total
1	Asfalto	365	470	835
2	Grama	480	474	954
3	Árvore	530	522	1052
4	Telhado	451	460	911

FONTE: O AUTOR, 2018.

3.3 MÉTODOS

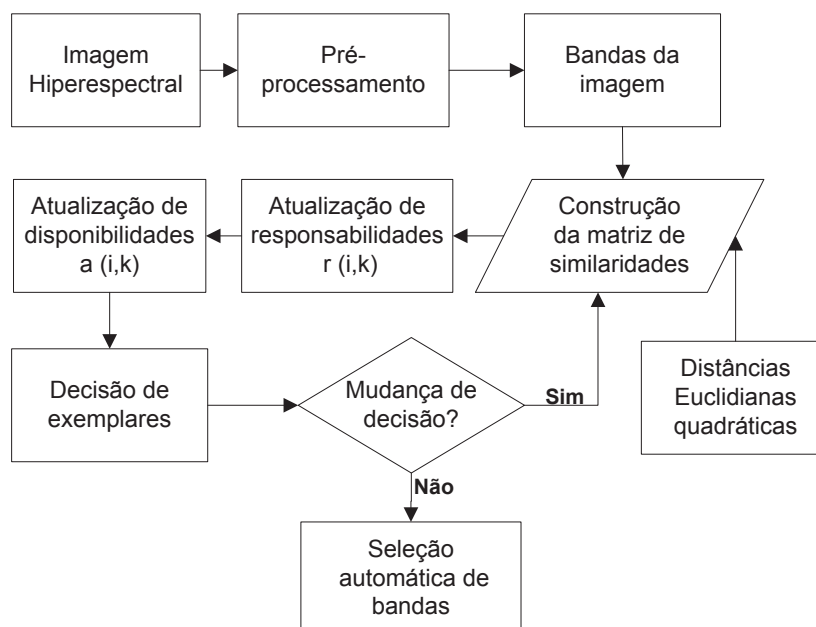
Nesta tese, é apresentada uma abordagem para *clusterização* e classificação de imagens hiperespectrais AVIRIS e HYDICE pela integração do algoritmo de *clusterização* Propagação de Afinidade baseado em Classes (PAC) e o classificador paramétrico Máxima Verossimilhança (MAXVER). Para efeitos de comparação e validação do método proposto, foram usados a PA original e a Análise de Componentes Principais (PCA). As etapas de implementação da metodologia proposta para seleção supervisionada e automática de bandas na imagem hiperespectral por respectivamente PAC-MAXVER e PA-MAXVER são apresentadas nos fluxogramas das Figuras 10 e 11.

FIGURA 10 - FLUXOGRAMA DA IMPLEMENTAÇÃO DA METODOLOGIA PROPOSTA PARA O MÉTODO PAC - MAXVER.



FONTE: O AUTOR, 2018.

FIGURA 11 - FLUXOGRAMA DA IMPLEMENTAÇÃO DA METODOLOGIA PROPOSTA PARA O MÉTODO PA - MAXVER.



FONTE: O AUTOR, 2018.

A implementação da metodologia do trabalho observou as seguintes etapas:

i) aquisição das imagens hiperespectrais; ii) pré-processamento dos dados coletados; iii) processamento propriamente dito; iv) pós-processamento da imagem classificada; e v) avaliação da acurácia e validação da classificação, explicadas nos tópicos seguintes:

3.3.1 Pré-processamento das imagens

Nesta etapa inicial da metodologia foi usado o ENVI 4.6.1 para remoção de bandas com informação redundante e ruído, tais como interferência do vapor de água na atmosfera, nas duas imagens hiperespectrais. Na sequência, foi usado o MultiSpec 3.4.1 na coleta de amostras de treinamento para seleção supervisionada de bandas das imagens através da *clusterização* por PA. Para terminar a etapa de pré-processamento, novas amostras para treinamento e teste foram coletadas nas imagens usando o ENVI 4.6.1. Essas amostras foram usadas no processo de classificação das imagens por MAXVER e pela integração PA-MAXVER e PCA-MAXVER, e posteriormente para avaliação da acurácia e validação dos resultados.

A imagem AVIRIS usada neste estudo é composta originalmente por 220 bandas espectrais, mas inclui regiões do espectro com ruídos existentes pela interferência do vapor de água na atmosfera. Estes ruídos podem causar inconveniências no momento da análise e processamento de dados (REFIANTI; MUTIARA; SYAMSUDDUHA, 2016), o que leva à necessidade de realizar exclusão e remoção preliminar das bandas com ruído (RAMAKRISHNAN; BHARTI, 2015).

A remoção das bandas foi feita com base na absorção de constituintes atmosféricos, portanto, as bandas com ruído foram identificadas das demais sem ruído. Após a identificação, essas bandas foram removidas das duas imagens mediante o procedimento já explicado nas seções anteriores.

3.3.2 Processamento

A etapa de processamento consistiu na extração de características da imagem para redução da dimensionalidade e seleção das bandas por *clusterização* através da PAC, PA, PCA e a classificação da imagem por MAXVER.

Foram realizados quatro experimentos nas duas imagens disponíveis para o estudo, nomeadamente AVIRIS e HYDICE. O Experimento I, designado PAC-MAXVER, faz a integração da Propagação de Afinidade baseada nas Classes (PAC) e a Máxima Verossimilhança (MAXVER), com duas fases (supervisionada e não supervisionada/automática) da PA.

O Experimento II também integra a PA e MAXVER, mas em uma única fase não supervisionada da PA. Este experimento foi designado PA-MAXVER. O Experimento III (PCA-MAXVER), combina o método de Análise de Componentes Principais (PCA) e o MAXVER; e finalmente o Experimento IV, o MAXVER, onde é realizada a classificação da imagem hiperespectral com todas as bandas, aplicando exclusivamente a Máxima Verossimilhança.

EXPERIMENTO I (PAC-MAXVER)

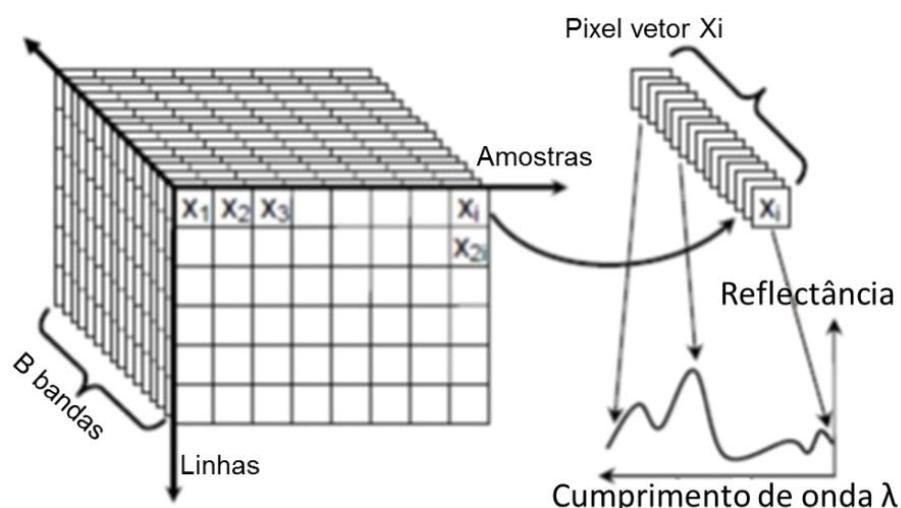
Para o método PAC-MAXVER, a seleção de bandas na imagem e redução de dimensionalidade por Propagação de Afinidade foi feita numa primeira fase supervisionada (*clusterização* baseada em amostras de treinamento por classes) e depois o processo foi não supervisionado (*clusterização* baseada em bandas), compreendendo as seguintes etapas:

Construção da Matriz de Similaridade

A matriz de similaridade é usada para agrupar pontos ou nódulos para formar *clusters*. Após o pré-processamento de dados, foram coletadas amostras de treinamento por classes para construção da matriz de similaridade. Para tal, foram consideradas duas abordagens: baseada nas classes (supervisionada) e baseada nas bandas (não supervisionada/automática).

Na Propagação de Afinidade supervisionada, baseada na classe, foram selecionadas várias amostras de treinamento com pixels vetores pertencentes a mesma classe, para todas as classes definidas nas imagens AVIRIS e HYDICE, conforme Figura 12. Na figura, é mostrada a estrutura de uma imagem hiperespectral, realçando o comportamento de amostras de pixels vetores nas diferentes bandas do espectro.

FIGURA 12 - ESTRUTURA DE UMA IMAGEM HIPERESPECTRAL



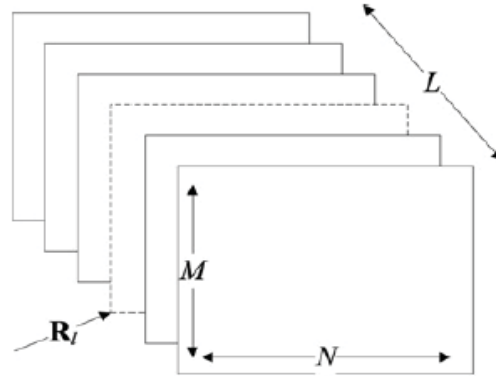
FONTE: TARABALKA; CHANUSSOT; BENEDIKTSSON, 2010.

Depois de coletadas as áreas de treinamento, todos os pixels vetores representando a mesma classe foram organizados pelas coordenadas espaciais linha *versus* coluna e os valores digitais de cada pixel foram extraídos usando uma rotina MATLAB desenvolvida no decurso dessa tese. Este procedimento (extração dos valores digitais dos pixels das amostras) foi realizado para todas as amostras de treinamento e no final gerado um ficheiro ascii (txt) com todas as amostras das classes representadas pelos valores digitais dos pixels, para determinação das similaridades ou afinidades por classe. Para o efeito foi aplicada a distância Euclidiana Quadrática Negativa entre os vários pares de pontos (valores digitais dos pixels) da referida classe, de acordo com a equação 14, segundo FREY e DUECK (2007).

As matrizes de similaridades obtidas por classe foram então usadas como dados de entrada no algoritmo de *clusterização* Propagação de Afinidade, uma por vez. Este procedimento foi repetido para todas as classes nas imagens AVIRIS e HYDICE.

Na fase não supervisionada, cada banda da imagem obtida através da *clusterização* supervisionada por PA foi considerada como um ponto (*data point*), como será explicado a seguir, com auxílio da Figura 13.

FIGURA 13 - REPRESENTAÇÃO DO CUBO DA IMAGEM HIPERESPECTRAL E CADA IMAGEM BANDA COMO PONTO.



FONTE: QIAN; YAO; JIA, 2009.

A partir da Figura 13, pode ser observado que uma imagem hiperespectral é um arranjo tridimensional com largura e comprimento correspondentes as dimensões espaciais e as bandas espectrais a terceira dimensão. Essas dimensões são neste trabalho, denotadas por M , N e L na sequência, e R é um cubo da imagem com cada banda $R_l \in \mathbb{R}^{M \times N}$ sendo uma matriz imagem. Para o presente trabalho, foram calculadas as médias das bandas das imagens, e cada imagem banda obtida após esse procedimento foi considerada um *data point* com dimensões $M \times N$. A partir da distância Euclidiana Quadrática Negativa sugerida por Frey e Dueck (2007), através da equação 14, foram achadas as diferenças entre todos pares de *data points* ou bandas espectrais existentes nas imagens. Obtidas todas as distâncias entre pares de pontos, isto é, similaridades, foram construídas duas matrizes de similaridades de 190×190 e de 162×162 , respectivamente para as imagens AVIRIS e HYDICE.

Cálculo de Exemplares / Evidências

O cálculo de exemplares ou evidências compreende o processo de propagação de mensagens de responsabilidade e disponibilidade no gráfico de fatores, conforme mostrado na Figura 5. Este processo inicializa com a matriz disponibilidade $a(i,k) = 0$ para a primeira iteração, e a matriz responsabilidade $r(i,k)$ inicializada como $S(i,k)$ ao substituir $a(i,k)$ na equação 15. A partir da segunda iteração, $a(i,k)$ não mais será igual a zero, e então as matrizes de responsabilidade e disponibilidade são atualizadas, através das equações 15 e 16.

Na fase supervisionada, todos os cálculos de evidência no algoritmo de *clusterização* Propagação de Afinidade baseado nas Classes foram realizados estabelecendo o parâmetro de preferência como o mínimo da diagonal da matriz de similaridade ($p = \min S(i,i)$), de modo a obter um número reduzido de *clusters*. Após experimentos realizados no decurso desta tese, foi escolhido o fator de amortecimento λ igual a 0,9. Este valor para λ , mostrou ser o mais adequado para atenuar as oscilações causadas durante o mecanismo de propagação de mensagens.

A fase não supervisionada (automática) foi a última e realizada separadamente, usando todas bandas representativas das classes, obtidas no decurso da fase supervisionada. Foi considerado para o cálculo das evidências, a mediana da diagonal da matriz de similaridades como parâmetro de preferência ($p = \text{median } S(i,i)$) porque de acordo com os experimentos, esse valor de preferência permite encontrar soluções de *clusterização* moderadas. O fator de amortecimento (*damping factor*) também foi estabelecido para λ igual a 0,9, como na fase anterior.

Atribuição de Exemplos ao Cluster

No sentido de atribuir exemplos ou centros de *clusters* a partir da matriz de similaridade, são considerados os seguintes procedimentos: primeiramente as matrizes de responsabilidade e de disponibilidade são criadas e atualizadas. Em seguida, os resultados da *clusterização* são determinados e os exemplos atribuídos através da equação 17.

Da matriz obtida pela equação 17, os exemplos podem ser atribuídos ou escolhidos facilmente fazendo a seleção dos valores positivos na diagonal principal dessa matriz. Esse processo de propagação de mensagens e atualização das responsabilidades e disponibilidades até a identificação do *cluster*, é repetido até que atinja um número especificado de iterações ou até que a estrutura do *cluster* se estabilize com um determinado número de iterações (DUECK, 2009).

Seleção de bandas representativas das classes na imagem

O passo final para a seleção de bandas da imagem baseado na Propagação de Afinidade é a descorrelação, realizada ao escolher apenas bandas

representativas de cada *cluster*. Nessas condições, os centros dos *clusters* são geralmente considerados as bandas preferíveis, visto serem altamente correlacionadas às restantes bandas pertencentes ao mesmo *cluster*, porque a similaridade dentro dos *clusters* é alta e a similaridade entre *clusters* é baixa. A remoção de bandas correlacionadas é um passo muito importante para redução de dimensionalidade de uma imagem hiperespectral. A seleção dos centros dos *clusters* como representativos das bandas que formam o *cluster* foi realizada considerando uma classe para cada *clusterização* por Propagação de Afinidade, e as bandas resultantes dessa seleção são representativas da referida classe.

Após obtenção das bandas representativas das classes, essas são combinadas para serem usadas na segunda etapa do método. Nesta etapa, diferentemente da primeira, a *clusterização* por PA não é supervisionada, e a matriz de similaridade é calculada a partir das distâncias Euclidianas Quadráticas Negativas entre os pares de bandas selecionadas na etapa anterior, usando a mesma métrica de similaridade proposta por Frey e Dueck (2007), dada pela equação 14. Porém, desta vez a mediana da similaridade foi considerada como parâmetro de preferência e fator de amortecimento foi de 0,9. As bandas obtidas por *clusterização* no final desse procedimento são usadas para integrar a Propagação de Afinidade e a Máxima Verossimilhança no processo de classificação.

EXPERIMENTO II (PA-MAXVER)

Construção da Matriz de Similaridade

Neste experimento foi aplicado o procedimento normal da Propagação de Afinidade, aplicando a métrica de similaridade dada pela equação 14, sobre as imagens hiperespectrais originais após pré-processamento. Seguindo o procedimento explicado anteriormente, cada banda da imagem foi considerada como um ponto de dimensão $M \times N$, e as similaridades entre as bandas da imagem foram calculadas diretamente entre cada par de bandas, através das distâncias Euclidianas Quadráticas Negativas, sugeridas por FREY e DUECK (2007). Com todos os valores das similaridades entre as bandas obtidos para as duas imagens, foram construídas as matrizes de similaridade de 190×190 e de 162×162 , respectivamente para as imagens AVIRIS e HYDICE.

Cálculo de Exemplos ou Evidências

Para o cálculo das evidências, inicializa-se a matriz de disponibilidades $a(i,k)$ = 0 para a primeira iteração, e ao substituí-las na equação 15, a matriz de responsabilidades fica, $r(i,k) = S(i,k)$. A partir da segunda iteração, $a(i,k)$ não mais será igual a zero, e então as matrizes de responsabilidade e disponibilidade são atualizadas, através das equações 15 e 16.

Neste experimento, as evidências foram calculadas considerando a mediana da diagonal da matriz de similaridade como parâmetro de preferência ($p = \text{median } S(i,i)$) com vista a permitir soluções de *clusterização* moderadas, e o fator de amortecimento λ igual a 0,9. Após a *clusterização*, as bandas representativas dos *clusters* formados, isto é, os centros dos *clusters*, foram selecionadas para usar no processo de classificação das imagens hiperespectrais por Máxima Verossimilhança.

Atribuição de Exemplos ao Cluster

Os exemplos dos *clusters* são atribuídos obedecendo os mesmos procedimentos do Experimento I.

Seleção de bandas representativas das classes na imagem

Neste passo final para a seleção de bandas aplicando a Propagação de Afinidade, é adotado o mesmo procedimento do experimento anterior, e as bandas representativas de cada *cluster* são selecionadas. A seleção dos centros dos *clusters* como representativos das bandas que formam o *cluster* foi realizada considerando os centros dos *clusters* obtidos numa única *clusterização* por Propagação de Afinidade, diferentemente do primeiro método em que a *clusterização* foi feita por classe, e as bandas obtidas usadas na classificação por Máxima Verossimilhança.

EXPERIMENTO III (PCA-MAXVER)

Para esse experimento, foi realizada a redução de dimensionalidade por Análise de Componentes Principais sobre as imagens hiperespectrais pré-processadas, com vista a extrair as bandas não correlacionadas, para subsequente classificação por Máxima Verossimilhança.

Para o cálculo das Componentes Principais sobre as duas imagens, AVIRIS e HYDICE, foram utilizados os seguidos procedimentos:

Primeiro foram coletadas M amostras de vetores de dimensão n , tendo como variáveis as colunas da imagem.

Em seguida, e para que o PCA funcione adequadamente, a imagem foi ajustada fazendo a subtração de cada um dos elementos dos vetores de dimensão da imagem pela sua respectiva média. Esse procedimento torna a média das colunas igual a zero e as variâncias unitárias, e evita que as colunas da imagem influenciem indevidamente nas Componentes Principais. A imagem ajustada pela média foi então usada para calcular a matriz de covariância da imagem, que é uma matriz quadrada $n \times n$, dada pelo produto entre a nova imagem ajustada e sua transposta.

A partir da matriz de covariância calculada, são calculados os autovalores e autovetores da mesma, e de seguida foi arranjada a matriz de Transformada de Hotelling. Essa matriz foi construída de modo que as linhas sejam formadas a partir dos autovetores da matriz de covariância, arranjados de modo que a primeira linha, o elemento (0,0), seja o autovetor correspondente ao maior autovalor, e assim sucessivamente até que a última linha corresponda ao menor autovalor. Essa transformação mapeia um vetor de dados a partir de um espaço original de p variáveis para um novo espaço de p variáveis não correlacionadas sobre o mesmo conjunto de dados. No entanto, nem todas as componentes principais precisam ser mantidas. Porém, foram mantidas somente as primeiras componentes principais, produzidas usando somente os primeiros autovetores, porque estas representam mais de 99,7% da informação contida (maior variância) nas bandas das imagens. Com esse procedimento, houve uma redução de dimensionalidade da imagem de 190 bandas após pré-processamento para 10 bandas (imagem AVIRIS), e de 162 bandas para 9 bandas (imagem HYDICE). Após a redução da dimensionalidade, a nova imagem obtida por PCA foi usada para classificação por MAXVER.

EXPERIMENTO IV (MAXVER)

Para realização do Experimento IV foram obedecidos os seguintes procedimentos:

Determinação dos tipos de cobertura de terra dentro da área de estudo;

Coleção dos pixels de treinamento para cada uma das classes desejadas;

Estimativa do vetor médio e da matriz de covariância de cada classe, a partir dos pixels de treinamento coletados;

Após o cálculo do vetor médio e da matriz covariância, cada pixel da imagem foi classificado para uma das classes desejadas ou rotulado como desconhecido.

Para a classificação por Máxima Verossimilhança, cada classe foi incluída numa região no espaço hiperespectral onde sua função discriminante fosse maior que a das outras classes.

3.3.3 Classificação da imagem

Depois da seleção de bandas e redução de dimensionalidade da imagem, o *software* ENVI 4.6.1 foi usado para realização da classificação da imagem. Para o efeito, o classificador paramétrico MAXVER foi adotado e aplicado da seguinte forma:

Em primeiro lugar e com auxílio das Figuras 8 e 9, os tipos de cobertura de terra ou classes dentro da área de estudo foram determinados para ambas as imagens, AVIRIS e HYDICE.

Seguindo a determinação das classes, as amostras de treinamento para cada uma das classes determinadas foram coletados nas duas imagens com base na informação da área de estudo ou verdade de campo dados nas Tabelas 1 e 2.

As amostras coletadas e treinadas foram então usadas para estimar o vetor médio e matriz de covariância de cada classe.

Finalmente, cada pixel nas imagens finais, obtidas pela combinação das bandas selecionadas por PAC, PA, PCA e as imagens inteiras pré-processadas, foi classificado por MAXVER numa das classes previamente definidas ou rotulado como desconhecido.

3.3.4 Avaliação da Acurácia e Validação dos Resultados

3.3.4.1 Avaliação da Acurácia da Classificação

A avaliação da acurácia foi realizada pela comparação do desempenho do método proposto com três outros métodos, incluindo:

- 1) a classificação da imagem hiperespectral por PA-MAXVER após a redução da dimensionalidade da imagem pela Propagação de Afinidade original;
- 2) a classificação da imagem hiperespectral por PCA-MAXVER após a redução da dimensionalidade da imagem, aplicando a análise de componentes principais;
- 3) a classificação da imagem hiperespectral original por MAXVER, usando todas as bandas da imagem.

Estas abordagens serão consideradas utilizando a acurácia do usuário (AU), acurácia do produtor (APr), acurácia global (AG), erros de omissão (EO), erros de inclusão (EI) e o coeficiente kappa (CK) da matriz de confusão, segundo as equações 22, 23, 24 e 25.

3.3.4.2 Validação dos Resultados da Classificação

A validação dos resultados da classificação pelo método proposto foi feita, considerando para os experimentos duas imagens hiperespectrais de dois ambientes completamente diferentes, nomeadamente a imagem AVIRIS para ambiente rural composto majoritariamente por diferentes tipos de vegetação, e a imagem HYDICE para ambiente urbano, composta na sua maioria por edifícios urbanos e vias de acesso. Para além das imagens representando os diferentes tipos de cobertura de terra e, obviamente níveis de complexidade diferentes, foi também usada nesta tese, para efeitos comparativos na classificação, a Propagação de Afinidade tradicional, a Análise por Componentes Principais e o método tradicional de classificação por Máxima Verossimilhança.

3.3.5 Pós – processamento da Imagem Classificada

Após a classificação, a imagem temática obtida pelo método proposto apresentou alguns pixels ou pontos isolados com efeitos aparentemente de ruídos ou erros na classificação. A presença desses erros/ruídos é minimizada nesta tese, realizando uma análise de contexto da classificação, aplicando o filtro da mediana.

Este filtro é muitas vezes usado para remover o ruído de uma imagem ou sinal, e essa redução de ruído é uma etapa típica de pré- ou pós-processamento que melhora os resultados da classificação segundo advoga CENTENO (2009).

A filtragem pela mediana foi escolhida porque além de remover o ruído, ela preserva as bordas das feições, melhorando deste modo a aparência visual da imagem classificada. Foi aplicada na filtragem, uma máscara bidimensional de tamanho 3×3 , cujos pixels foram classificados na ordem dos seus níveis de cinza, e o valor mediano do grupo foi guardado para substituir os valores dos pixels ruidosos, de acordo com a equação 21.

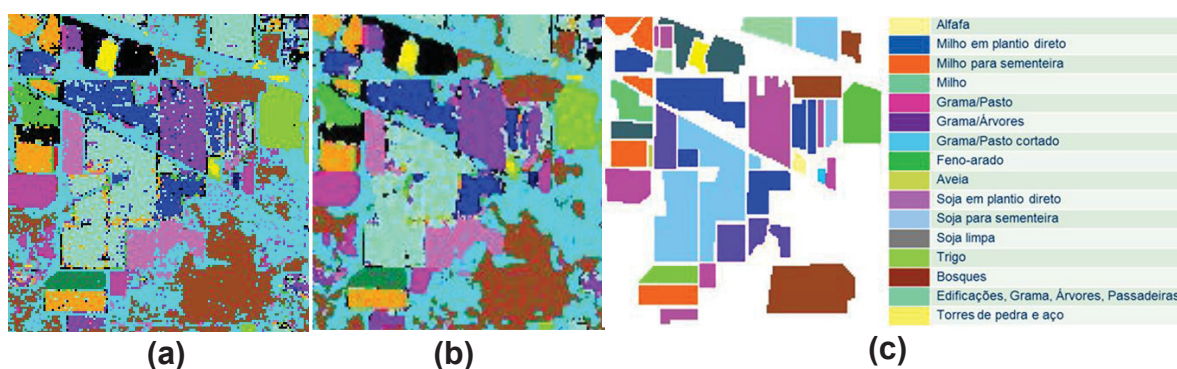
CAPÍTULO 4. APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Os resultados obtidos dos Experimentos I, II, III e IV realizados nesta tese, para a zona rural, coletada pelo sensor AVIRIS e para a zona urbana, coletada pelo sensor HYDICE são apresentados nas Figuras 14 – 23 e Tabelas 4 – 13, a seguir:

4.1 EXPERIMENTOS COM A IMAGEM AVIRIS

4.1.1 Experimento I – Método PAC – MAXVER

FIGURA 14 - IMAGENS CLASSIFICADAS SEM APLICAÇÃO DO FILTRO DE MEDIANA (A), COM APLICAÇÃO DO FILTRO DE MEDIANA (B) E VERDADE DE CAMPO PARA O EXPERIMENTO I.



FONTE: O AUTOR, 2018.

TABELA 4 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO I COM A IMAGEM AVIRIS.

PAC-MAXVER Classes	Acurácia (%)			
	APr	EO	AU	EI
Alfafa	92,86	7,14	100	0,00
Milho-plantio-direto	43,33	56,67	81,25	18,75
Milho-sementeira	100,0	0,00	100,0	0,00
Grama-pasto	100,0	0,00	100,0	0,00
Grama-árvore	100,0	0,00	92,59	7,41
Grama-pasto cortado	100,0	0,00	100,0	0,00
Feno arado	100,0	0,00	97,92	2,08
Aveia	100,0	0,00	100,0	0,00
Soja-plantio-direto	81,25	18,75	100,0	0,00
Soja-sementeira	100,0	0,00	54,84	45,16
Soja	100,0	0,00	100,0	0,00
Trigo	100,0	0,00	100,0	0,00
Bosque	100,0	0,00	92,31	7,69
Edif-grama-passadeiras	87,10	12,90	100,0	0,00
Torres	100,0	0,00	100,0	0,00
Milho	100,0	0,00	96,55	3,45
AG	94,15			
CK	0,94			

FONTE: O AUTOR, 2018.

A Figura 14 e a Tabela 4 apresentam os resultados do Experimento I, realizado pelo método de classificação PAC – MAXVER proposto, usando a imagem hiperespectral AVIRIS. Na Figura 14 (a) a esquerda, é mostrado o resultado da classificação da imagem sem uso do filtro da mediana, enquanto que na Figura 14 (b) a imagem classificada é mostrada após ter sido aplicado o filtro da mediana.

Para o processo de classificação da imagem foram consideradas as dezesseis (16) classes nela contidas, conforme ilustrado na Tabela 4. De acordo com os resultados, o método PAC-MAXVER apresentou uma separação das classes com uma acurácia global de 94,15% e coeficiente kappa de 0,94. Esse valor de acurácia obtido ao empregar este método explica que 94,15% dos pixels na imagem foram corretamente classificados, sendo apenas 5,85% destes classificados com erros. O método PAC-MAXVER foi aplicado após ter resultado numa redução de dimensionalidade de 190 para 9 bandas, e classificou as 16 classes definidas na imagem original.

Os resultados da Tabela 4 ilustram que o PAC-MAXVER resultou em uma acurácia do produtor de 100% para 12 classes, o que significa que somente 4 classes, nomeadamente alfafa, milho-plantio-direto, soja-plantio-direto e edif-grama-passadeiras apresentaram erros na classificação para o produtor.

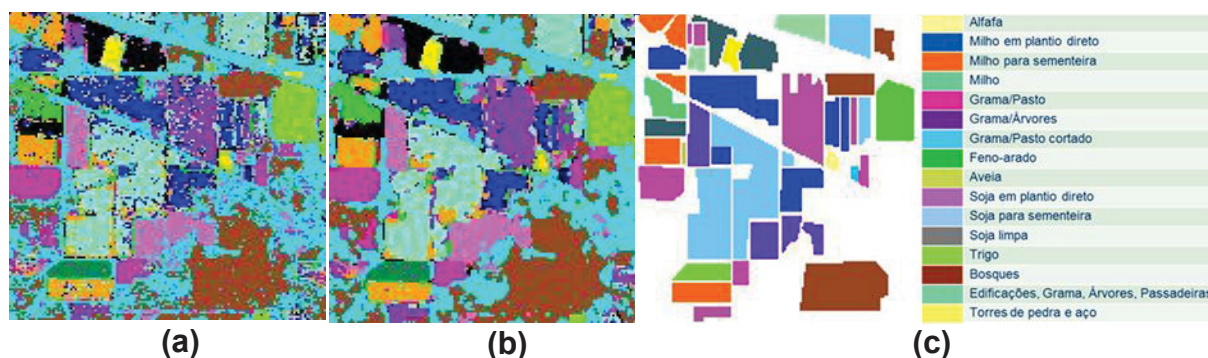
Durante o processo de classificação, a classe alfafa foi omissa em 7,14%, milho-plantio-direto foi omissa em 56,67% e atribuída erradamente por inclusão em 18,75%, soja-plantio-direto omissa em 18,75% e edif-grama-passadeiras em 12,9%. Das classes com erros na acurácia do produtor, a classe milho-plantio-direto foi aquela que registou menor acurácia, isto é, maior erro, por volta de 43%.

Fazendo uma análise visual da imagem classificada na Figura 14(b), observa-se que após filtragem da imagem, a classe soja-plantio-direto na zona central aparece melhorada, sendo maior parte dos pixels isolados e/ou ruídos que se apresentam na Figura 14(a) foram removidos.

Neste experimento, o valor do coeficiente kappa (0,94) foi acima de 0,9, o que de acordo com Cohen (1960) e Landis e Koch (1977), sugere que o método PAC-MAXVER apresentou uma concordância quase perfeita entre as classes da imagem avaliada e boa acurácia.

4.1.2 Experimento II – Método PA – MAXVER

FIGURA 15 - IMAGENS CLASSIFICADAS SEM APLICAÇÃO DO FILTRO DA MEDIANA (A), COM APLICAÇÃO DO FILTRO DA MEDIANA (B) E VERDADE DE CAMPO PARA O EXPERIMENTO II.



FONTE: O AUTOR, 2018.

TABELA 5 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO II COM A IMAGEM AVIRIS.

PA-MAXVER	Acurácia (%)			
Classes	APr	EO	AU	EI
Alfafa	87,50	12,50	100	0,00
Milho-plantio-direto	48,28	51,72	100	0,00
Milho-sementeira	100,0	0,00	82,98	17,02
Grama-pasto	100,0	0,00	85,71	14,29
Grama-árvore	90,91	9,09	83,33	16,67
Grama-pasto cortado	100,0	0,00	100	0,00
Feno arado	100,0	0,00	95,65	4,35
Aveia	100,0	0,00	100,0	0,00
Soja-plantio-direto	92,86	7,14	86,67	13,33
Soja-sementeira	100,0	0,00	85,71	14,29
Soja	90,24	9,76	100,0	0,00
Trigo	100,0	0,00	100,0	0,00
Bosques	100,0	0,00	93,94	6,06
Edif-grama-passadeiras	92,59	7,41	96,15	3,85
Torres	100,0	0,00	100,0	0,00
Milho	96,88	3,13	86,11	13,89
AG	93,61			
CK	0,93			

FONTE: O AUTOR, 2018.

Neste experimento foi usada a imagem hiperespectral AVIRIS e consideradas as dezesseis (16) classes para a classificação com o método PA-MAXVER. De acordo com os resultados da Tabela 5 e Figura 15, este método apresentou uma separação das classes com uma acurácia global de 93,61% e coeficiente kappa de 0,93. Este valor de acurácia obtido ao empregar este método, explica que 93,61% dos pixels na imagem AVIRIS foram corretamente classificados e 6,39% foram classificados com erros.

Para aplicar o método, foram usadas 14 bandas após redução de dimensionalidade, com as quais as 16 classes presentes na imagem foram classificadas. A Tabela 5 ilustra que o PA-MAXVER resultou em uma acurácia do produtor de 100% para 9 classes, o que significa que as restantes 7 classes, nomeadamente alfafa, milho-plantio-direto, grama-árvore, soja-plantio-direto, soja, edif-grama-passadeiras e milho apresentaram erros na classificação para o produtor.

Durante o processo de classificação, a classe alfafa foi omissa em 12,50%, milho-plantio-direto foi omissa em 51,72%, grama-árvore foi omissa em 9,09% e atribuída por inclusão em 16,67%, soja-plantio-direto omissa em 7,14% e atribuída por inclusão em 13,33%, soja foi omissa em 9,76%, edif-grama-passadeiras foi omissa em 7,41% e atribuída por inclusão em 3,85% e milho foi omissa em 3,13% e atribuída por inclusão em 13,89%. Das classes com erros na acurácia do produtor, a classe milho-plantio-direto foi aquela que registou menor acurácia (por conseguinte a mais omitida), estando por volta de 48%, correspondente a 52% de omissão.

Analizando visualmente a imagem classificada na Figura 15(b), observa-se que junto a classe soja para sementeira na parte inferior esquerda apresenta erros de classificação ou ruídos que o filtro da mediana não conseguiu remover.

De acordo com Cohen (1960) e Landis e Koch (1977), o valor do coeficiente Kappa de 0,93, acima de 0,9 sugere que o método PA – MAXVER apresenta uma concordância quase perfeita entre as classes da imagem avaliada e uma boa acurácia.

4.1.3 Experimento III – Método PCA – MAXVER

FIGURA 16 - IMAGEM CLASSIFICADA USANDO O EXPERIMENTO III E VERDADE DE CAMPO.



FONTE: O AUTOR, 2018.

TABELA 6 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO III COM A IMAGEM AVIRIS.

PCA-MAXVER	Acurácia (%)			
Classes	APr	EO	AU	EI
Alfafa	88,24	11,76	100	0,00
Milho-plantio-direto	46,15	53,85	92,31	7,69
Milho-sementeira	89,47	10,53	100,0	0,00
Grama-pasto	100,0	0,00	100,0	0,00
Grama-árvore	100,0	0,00	100,0	0,00
Grama-pasto cortado	100,0	0,00	100,0	0,00
Feno arado	100,0	0,00	100,0	0,00
Aveia	100,0	0,00	100,0	0,00
Soja-plantio-direto	80,0	20,0	100,0	0,00
Soja-sementeira	100,0	0,00	70,59	29,41
Soja	100,0	0,00	96,97	3,03
Trigo	100,0	0,00	100,0	0,00
Bosques	100,0	0,00	83,78	16,22
Edif-grama-passadeiras	81,82	18,18	81,82	18,18
Torres	100,0	0,00	100,0	0,00
Milho	100,0	0,00	80,0	20,0
AG	93,24			
CK	0,93			

FONTE: O AUTOR, 2018.

No Experimento III, pelo método PCA – MAXVER realizado usando a imagem hiperespectral AVIRIS, foram consideradas as dezesseis (16) classes. Os resultados obtidos no experimento são apresentados na Figura 16 e Tabela 6.

De acordo com os resultados da tabela, este método apresentou uma separação das classes, com uma acurácia global de 93,24% e coeficiente kappa de 0,93. Este valor de acurácia global obtido ao empregar este método explica que um total de 93,24% dos pixels na imagem foram classificados corretamente e apenas 6,76% dos pixels foram classificados erradamente.

O método PCA - MAXVER foi aplicado após uma redução de dimensionalidade da imagem original de 190 para 10 bandas, e classificou as 16 classes definidas na imagem. Os resultados na Tabela 6 ilustram que o PCA - MAXVER resultou em uma acurácia do produtor de 100% para 11 classes, o que significa que somente 5 classes, nomeadamente alfafa, milho-plantio-direto, milho-sementeira, soja-plantio-direto e edif-grama-passadeiras apresentaram erros na classificação para o produtor. Essas cinco classes, na mesma ordem em que se apresentam, foram omissas em respectivamente 11,76%, 53,85%, 10,53%, 20% e 18,18%. Os erros de inclusão neste método foram de 7,69%, 29,41%, 3,03%, 16,22%, 18,18% e 20% respectivamente para as classes milho-plantio-direto, soja-sementeira, soja, bosques, edif-grama-passadeiras e milho. Analisando visualmente

a imagem, observam-se alguns pixels isolados e ruidosos sobre as classes soja para sementeira e soja em plantio direto na zona central da imagem.

O valor de coeficiente kappa de 0,93 apresentado pelo PCA – MAXVER foi maior que 0,9, o que de acordo com Cohen (1960) e Landis e Koch (1977), sugere uma concordância quase perfeita entre as classes da imagem avaliada e boa acurácia.

4.1.4 Experimento IV – Método MAXVER

FIGURA 17 - IMAGEM CLASSIFICADA USANDO O EXPERIMENTO IV E VERDADE DE CAMPO.



FONTE: O AUTOR, 2018.

TABELA 7 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO IV COM A IMAGEM AVIRIS.

MAXVER Classes	Acurácia (%)			
	APr	EO	AU	EI
Milho-plantio-direto	57,10	42,90	99,53	0,47
Milho-sementeira	73,71	26,29	100,0	0,00
Grama-pasto	97,20	2,80	100,0	0,00
Grama-árvore	98,66	1,34	96,08	3,92
Feno arado	100,0	0,00	100,0	0,00
Soja-plantio-direto	94,74	5,26	97,96	2,04
Soja-sementeira	100,0	0,00	50,32	49,68
Soja	90,18	9,82	100,0	0,00
Bosques	100,0	0,00	95,51	4,49
Edif-grama-passadeiras	88,35	11,65	100,0	0,00
OA	85,42			
KC	0,83			

FONTE: O AUTOR, 2018.

No Experimento IV realizado, foi usada a imagem hiperespectral AVIRIS, e consideradas as dezesseis classes existentes na imagem. Para o processo de classificação foi aplicado o método MAXVER, e os resultados obtidos são apresentados na Figura 17 e Tabela 7.

De acordo com os resultados da Tabela 7, o método MAXVER separou as classes com uma acurácia global de 85,42% e coeficiente kappa de 0,83. Este valor

de acurácia obtido ao empregar este método explica que 85,42% dos pixels na imagem foram corretamente classificados, sendo os restantes 14,58% dos pixels classificados com erros.

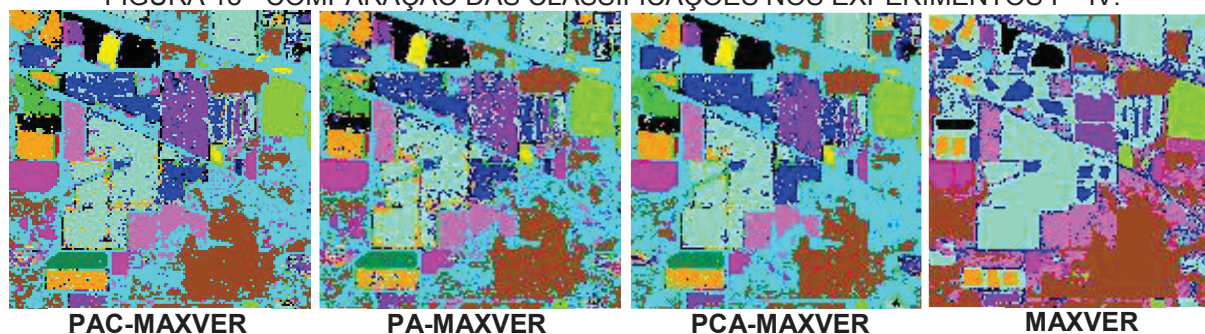
O método MAXVER foi aplicado com a imagem AVIRIS após pré-processamento, considerando das 220 bandas da imagem original, somente 190 bandas. Foram classificadas apenas 10 classes das 16 existentes na imagem. Esta limitação na classificação das classes é atribuída ao facto de as classes não classificadas, neste caso, alfafa, milho, grama-pasto cortado, aveia, trigo e torres, apresentarem um pequeno número de amostras de treinamento em um espaço de características de alta dimensionalidade, estando de acordo com LANDGREBE (2003) e CAMPS-VALLS et al. (2014).

Das 10 classes classificadas pelo método MAXVER, somente 3 (feno arado, soja-sementeira e bosques) foram classificadas corretamente, estando 7 com erros de classificação. Durante a classificação, as classes foram omissas da seguinte forma: milho-plantio-direto (42,90%), milho-sementeira (26,29%), grama-pasto (2,80%), grama-árvore (1,34%), soja-plantio-direto (5,26%), soja (9,82%) e edific-grama-passadeiras (11,65%). Neste mesmo processo, foram verificados erros de inclusão, onde foram atribuídas erradamente classes para milho-plantio-direto em 0,47%, grama-árvore em 3,92% e finalmente 2,04% para soja-plantio-direto.

O valor do coeficiente kappa obtido no método MAXVER foi de 0,83. Este valor está a baixo de 0,9, o que de acordo com Cohen (1960) e Landis e Koch (1977), sugere que o método MAXVER apresenta uma forte concordância entre as classes.

4.1.5 Análise Conjunta dos Resultados Obtidos pelos Métodos I – IV, Usando a Imagem Hiperespectral AVIRIS

FIGURA 18 - COMPARAÇÃO DAS CLASSIFICAÇÕES NOS EXPERIMENTOS I – IV.



FONTE: O AUTOR, 2018.

TABELA 8 - RESUMO DA ACURÁCIA DOS EXPERIMENTOS I – IV COM IMAGEM AVIRIS.

	PAC-MAXVER				PA-MAXVER				PCA-MAXVER				MAXVER			
Acur (%)	APr	EO	AU	EI	APr	EO	AU	EI	APr	EO	AU	EI	APr	EO	AU	EI
1	92.9	7.1	100	0	100	0	100	0	88.2	11.8	100	0	Não classificada			
2	43.3	56.7	81.3	18.8	48	52	63.2	36.8	46.2	53.9	92.3	7.7	57.1	42.9	99.5	0.5
3	100	0	100	0	91.4	8.6	100	0	89.5	10.5	100	0	73.7	26.3	100	0
4	100	0	96.6	3.5	100	0	73.7	26.3	100	0	80	20	Não classificada			
5	100	0	100	0	100	0	100	0	100	0	100	0	97.2	2.8	100	0
6	100	0	92.6	7.4	93.1	6.9	93.1	6.9	100	0	100	0	98.7	1.3	96.1	3.9
7	100	0	100	0	100	0	100	0	100	0	100	0	Não classificada			
8	100	0	97.9	2.1	100	0	100	0	100	0	100	0	100	0	100	0
9	100	0	100	0	100	0	83.3	16.7	100	0	100	0	Não classificada			
10	81.3	18.8	100	0	61.1	38.9	91.7	8.3	80	20	100	0	94.7	5.3	98	2
11	100	0	54.8	45.2	100	0	66.7	33.3	100	0	70.6	29.4	100	0	50.3	49.7
12	100	0	100	0	94.3	5.7	89.2	10.8	100	0	97	3	90.2	9.8	100	0
13	100	0	100	0	100	0	100	0	100	0	100	0	Não classificada			
14	100	0	92.3	7.7	100	0	91.8	8.2	100	0	83.8	16.2	100	0	95.5	4.5
15	87.1	12.9	100	0	84.4	15.6	96.4	3.6	81.8	18.2	81.8	18.2	88.4	11.7	100	0
16	100	0	100	0	100	0	100	0	100	0	100	0	Não classificada			
OA	94.15				92.20				93.24				85.42			
KC	0.94				0.92				0.93				0.83			

FONTE: O AUTOR, 2018.

1-Alfafa; 2-Milho-plantio-direto; 3-Milho-sementeira; 4-Milho; 5-Grama-pasto; 6-Grama-árvore; 7-Grama-pasto cortado; 8-Feno arado; 9-Aveia; 10-Soja-plantio-direto; 11-Soja-sementeira; 12-Soja; 13-Trigo; 14-Bosques; 15-Edifício-grama-passadeiras; 16-Torres

A Figura 18 e a Tabela 8 mostram o resumo dos experimentos realizados usando a imagem hiperespectral AVIRIS. Para o efeito, foram consideradas as dezesseis (16) classes presentes na imagem. Para o estudo foi proposto o método de classificação PAC-MAXVER, e para efeitos comparativos os métodos PA-MAXVER, PCA-MAXVER e MAXVER.

De acordo com os resultados apresentados na Tabela 8, o método PAC-MAXVER apresenta a melhor separação das classes, com uma acurácia global de 94,15% e coeficiente kappa de 0,94.

A classificação pelo método PA-MAXVER foi a segunda melhor, com uma acurácia global de 93,61% e coeficiente kappa de 0,93. A classificação usando o método PCA-MAXVER esteve ligeiramente a baixo dos dois primeiros métodos, mas muito acima do MAXVER e resultou numa acurácia global de 93,24% e coeficiente kappa de 0,93.

O método que apresentou menor valor ao separar as classes foi o tradicional MAXVER, com uma acurácia global de apenas 85,42% e coeficiente kappa de 0,83.

Os valores das acurácias obtidas ao empregar estes métodos explicam que 94,15%, 93,61, 93,24% e 85,42% dos pixels na imagem foram corretamente classificados, e respectivamente 5,85%, 6,39%, 6,75% e 14,58% dos pixels foram classificados com erros.

O método MAXVER foi aplicado com a imagem original, considerando todas 190 bandas, e classificou somente 10 classes, contrariamente a PAC-MAXVER, PA-MAXVER e PCA-MAXVER que classificaram as 16 classes. Esta limitação na classificação das classes da imagem é atribuída ao fato de as classes não classificadas (alfafa, milho, grama-pasto cortado, aveia, trigo e torres) apresentarem pequeno número de amostras de treinamento em um espaço de características de alta dimensionalidade, estando de acordo com LANDGREBE (2003) e CAMPS-VALLS et al. (2014).

Os resultados obtidos ilustram que o PAC-MAXVER resultou em uma acurácia do produtor de 100% para 12 classes, o que significa que somente 4 classes, nomeadamente alfafa, milho-plantio-direto, soja-plantio-direto e edif-grama-passadeiras apresentaram erros na classificação para o produtor.

Durante o processo de classificação, a classe alfafa foi omitida em 7,14%, milho-plantio-direto foi omitida em 56,67% e atribuída por inclusão em 18,75%, soja-plantio-direto omitida em 18,75% e edif-grama-passadeiras em 12,9%. Das classes com erros na acurácia do produtor, a classe milho-plantio-direto foi aquela que registou menor acurácia (maior erro) nos três métodos, sendo 43%, 46% e 57% respectivamente para PAC-MAXVER, PCA-MAXVER e MAXVER.

A melhor acurácia na classificação obtida pelo método baseado na Propagação de Afinidade por Classe (PAC-MAXVER) provou que esta, tal como outros métodos discriminativos não são influenciados negativamente pelo número limitado de amostras de treinamento como é atestado pela literatura (PENG et al., 2016).

No que se refere a acurácia do produtor obtida no método PCA-MAXVER, foi em 100% para as mesmas classes que o PAC-MAXVER, diferindo apenas em uma classe (milho-sementeira), o que dá a este método 11 classes corretamente classificadas, sendo 5 com erros de classificação. As cinco classes, alfafa, milho-plantio-direto, milho-sementeira, soja-plantio-direto e edif-grama-passadeiras foram omitidas em respectivamente 11,76%, 53,85%, 10,53%, 20% e 18,18%. Os erros de

inclusão neste método foram de 7,69% e 18,18% respectivamente para as classes milho-plantio-direto e edifício-grama-passadeiras.

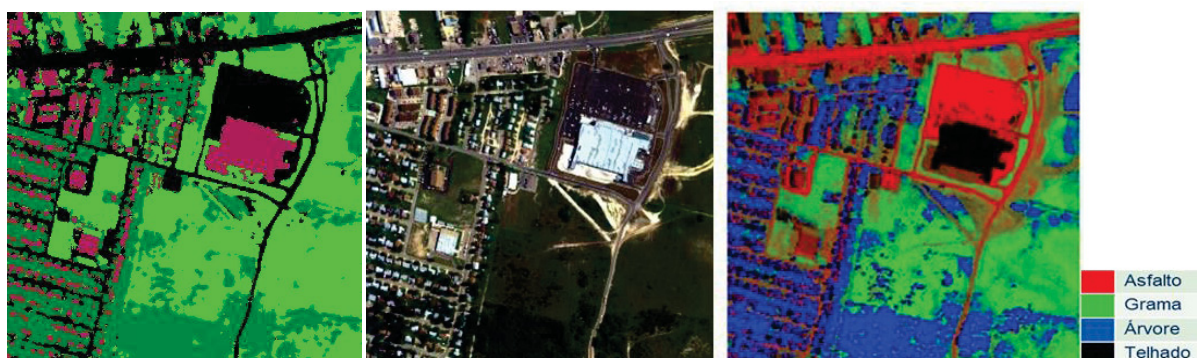
Das 10 classes classificadas pelo terceiro método, o MAXVER, somente 3 (feno arado, soja-sementeira e bosques) foram classificadas corretamente, estando 7 com erros de classificação. Durante a classificação, as classes foram omitidas da seguinte forma: milho-plantio-direto (42,90%), milho-sementeira (26,29%), grama-pasto (2,80%), grama-árvore (1,34%), soja-plantio-direto (5,26%), soja (9,82%) e edif-grama-passadeiras (11,65%). Neste mesmo processo, foram verificados erros de inclusão, onde foram atribuídas erradamente classes para milho-plantio-direto em 0,47%, grama-árvore em 3,92% e finalmente 2,04% para soja-plantio-direto.

Os valores dos coeficientes kappa para os Experimentos I, II, III e IV foram respectivamente 0,94, 0,93, 0,93 e 0,83. De acordo com Cohen (1960) e Landis e Koch (1977), esses valores sugerem que os métodos PAC-MAXVER, PA-MAXVER e PCA-MAXVER que possuem o coeficiente Kappa acima de 0,9 apresentam uma concordância quase perfeita entre as classes das imagens avaliadas e boa acurácia. O método MAXVER com kappa igual a 0,83 apresenta uma forte concordância entre as classes.

4.2 EXPERIMENTOS COM A IMAGEM HYDICE

4.2.1 Experimento I – Método PAC – MAXVER

FIGURA 19 - IMAGEM CLASSIFICADA COM APLICAÇÃO DO FILTRO DA MEDIANA, IMAGEM ORIGINAL E VERDADE DE CAMPO.



FONTE: O AUTOR, 2018.

TABELA 9 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO I COM A IMAGEM HYDICE.

PAC-MAXVER	Acurácia (%)			
Classes	APr	EO	AU	EI
Asfalto	97,54	2,46	99,65	0,35
Árvore	99,38	0,62	85,06	14,94
Gramma	92,60	7,40	98,55	1,45
Telhado	97,56	2,44	98,07	1,93
AG	96,47			
CK	0,95			

FONTE: O AUTOR, 2018.

No Experimento I realizado pelo método PAC – MAXVER proposto, foi usada a imagem hiperespectral HYDICE, e considerado um total de quatro (04) classes. De acordo com os resultados apresentados na Figura 19 e Tabela 9, a classificação baseada no método PAC-MAXVER resultou numa acurácia global de 96,47% e coeficiente kappa de 0,95. Esta acurácia obtida explica que 96,47% dos pixels na imagem foram corretamente classificados e que 3,53% dos pixels foram classificados com erros.

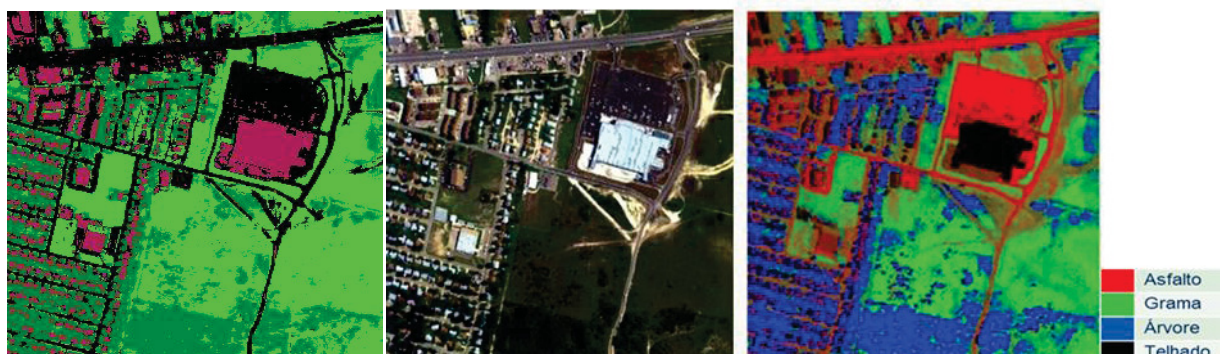
O método proposto foi aplicado com uma redução de dimensionalidade da imagem pré-processada de 162 bandas para 10 bandas, e classificou maioria das feições das 4 classes existentes.

Os resultados mostram também que o método PAC-MAXVER resultou numa acurácia do produtor de 97,54% para asfalto, 99,38% para árvore, 92,60% para grama e 97,56% para telhado, o que significa que os erros cometidos nas classes asfalto, grama e telhado foram abaixo de 2% e somente a classe árvore apresentou um erro de inclusão por volta dos 15%. Durante o processo de classificação, a classe asfalto foi omissa em 2,46%, árvore foi omissa em 0,62% e atribuída por inclusão em 14,94%, grama omitida em 7,40% e telhado em 2,44%.

O valor do coeficiente kappa para o Experimento I baseado no PAC – MAXVER foi maior que 0,9 (0,95), o que de acordo com Cohen (1960) e Landis e Koch (1977), sugere uma concordância quase perfeita entre as classes da imagem avaliada e boa acurácia.

4.2.2 Experimento II – Método PA – MAXVER

FIGURA 20 - IMAGEM CLASSIFICADA COM APLICAÇÃO DO FILTRO DA MEDIANA PARA O EXPERIMENTO II, IMAGEM ORIGINAL E VERDADE DE CAMPO.



FONTE: O AUTOR, 2018.

TABELA 10 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO II COM A IMAGEM HYDICE.

PA-MAXVER Classes	Acurácia (%)			
	APr	EO	AU	EI
Asfalto	95,58	4,42	99,34	0,66
Árvore	98,57	1,43	86,73	13,27
Gramas	92,13	7,87	96,48	3,52
Telhado	97,10	2,90	89,94	10,06
AG	95,26			
CK	0,93			

FONTE: O AUTOR, 2018.

Neste experimento foi usada a imagem hiperespectral HYDICE e foi aplicado sobre ela, o método de classificação PA-MAXVER, considerando as quatro (4) classes existentes na imagem.

De acordo com os resultados apresentados na Figura 20 e Tabela 10, o método PA-MAXVER apresentou separação entre as classes, com uma acurácia global de 95,26% e coeficiente kappa de 0,93. Este valor da acurácia obtido ao empregar este método explica que 95,26% dos pixels na imagem foram corretamente classificados, e 4,74% dos pixels foram classificados com erros. O método PA-MAXVER foi aplicado com redução de dimensionalidade de 162 bandas da imagem pré-processada para 9 bandas, e classificou as quatro classes.

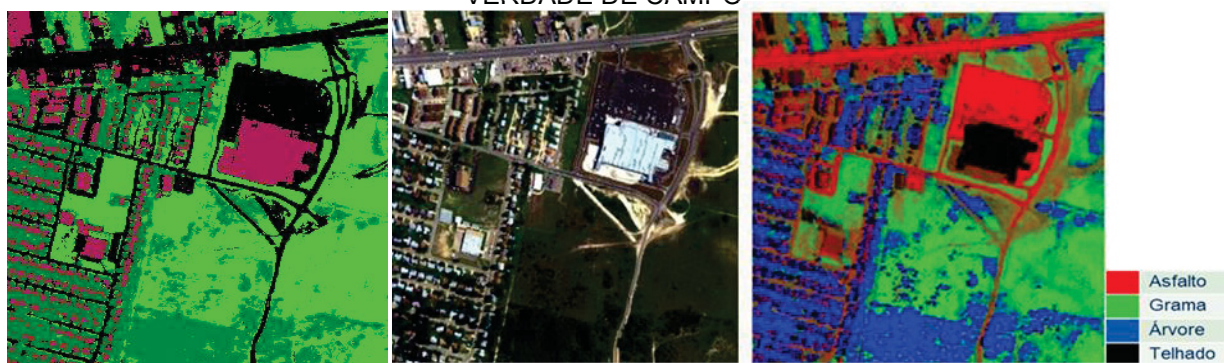
Os resultados na tabela mostram também que o PA-MAXVER resultou numa acurácia do produtor de 95,58% para asfalto, 98,57% para árvore, 92,13% para grama e 97,10% para telhado, o que significa que os erros cometidos por inclusão nas classes asfalto, árvore, grama e telhado foram respectivamente 0,66%, 13,27%,

3,52% e 10,06%. Durante o processo de classificação, a classe asfalto foi omitida em 4,42%, árvore foi omitida em 1,43%, grama omitida em 7,87% e telhado em 2,90%.

O valor do coeficiente kappa para esse experimento foi de 0,93, valor maior que 0,9. De acordo com Cohen (1960) e Landis e Koch (1977), esse valor sugere que o método PA-MAXVER apresenta uma concordância quase perfeita entre as classes das imagens avaliadas e boa acurácia.

4.2.3 Experimento III – Método PCA – MAXVER

FIGURA 21 - IMAGEM CLASSIFICADA PARA O EXPERIMENTO III, IMAGEM ORIGINAL E VERDADE DE CAMPO



FONTE: O AUTOR, 2018.

TABELA 11 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO III COM A IMAGEM HYDICE.

PCA-MAXVER	Acurácia (%)			
Classes	APr	EO	AU	EI
Ásfalto	97,04	2,96	99,87	0,13
Árvore	98,94	1,06	88,97	11,03
Grama	93,92	6,08	96,86	3,14
Telhado	99,07	0,93	95,74	4,26
AG	96,69			
CK	0,95			

FONTE: O AUTOR, 2018.

No Experimento III foi usado o método PCA-MAXVER com a imagem hiperespectral HYDICE e foram consideradas as quatro (04) classes na imagem. De acordo com os resultados apresentados na Figura 21 e Tabela 11, este método apresentou uma separação entre as classes, com uma acurácia global de 96,69% e coeficiente kappa de 0,95.

Os resultados mostram que o valor da acurácia obtida ao empregar o método PCA-MAXVER explica que 96,69% dos pixels na imagem foram corretamente

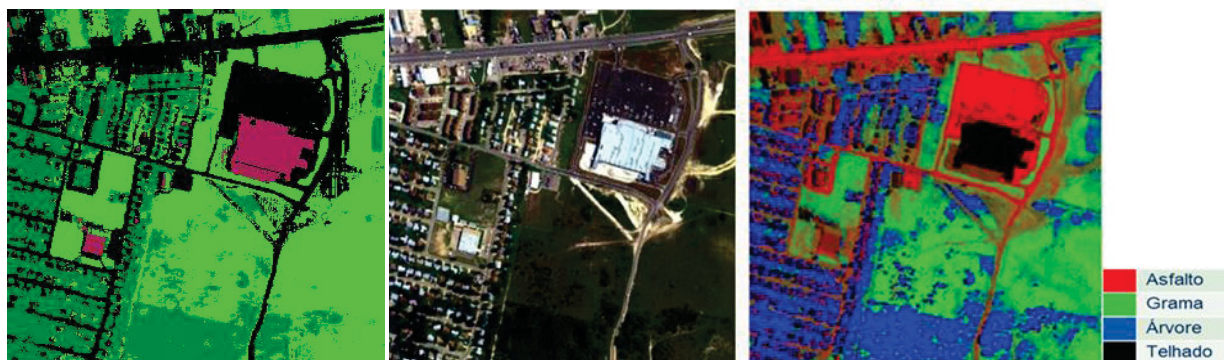
classificados, e 3,31% dos pixels foram classificados com erros. Este método foi aplicado após uma redução de dimensionalidade da imagem pré-processada de 162 bandas para 9 bandas, e classificou as 4 classes existentes na imagem.

O método PCA-MAXVER resultou numa acurácia do produtor de 97,04% para asfalto, 98,94% para árvore, 93,92% para grama e 99,07% para telhado, sendo os erros cometidos por inclusão para as mesmas classes e na mesma ordem foram respetivamente 0,13%, 11,03%, 3,14% e 4,26%. Durante o processo de classificação, a classe asfalto foi omitida em 2,96%, árvore foi omitida em 1,06%, grama omitida em 6,08% e telhado em 0,93%. As acurácias para o usuário para as classes em estudo estiveram todas acima de 95%.

O valor do coeficiente kappa para o método PCA-MAXVER foi acima de 0,9 (0,95), o que de acordo com Cohen (1960) e Landis e Koch (1977), significa que a classificação realizada pelo referido método apresentou uma concordância quase perfeita entre as classes das imagens avaliadas e boa acurácia.

4.2.4 Experimento IV – Método – MAXVER

FIGURA 22 - IMAGEM CLASSIFICADA PARA O EXPERIMENTO IV, IMAGEM ORIGINAL E VERDADE DE CAMPO.



FONTE: O AUTOR, 2018.

TABELA 12 - AVALIAÇÃO DA ACURÁCIA PARA O EXPERIMENTO IV COM A IMAGEM HYDICE.

MAXVER	Acurácia (%)			
	APr	EO	AU	EI
Asfalto	100	0,00	99,90	0,10
Árvore	0,00	0,00	0,00	100
Grama	0,46	99,54	100	0,00
Telhado	0,00	100	0,00	0,00
AG	81,50			
CK	0,45			

FONTE: O AUTOR, 2018.

Na Figura 22 e Tabela 12 são apresentados os resultados do Experimento IV, realizado com o tradicional MAXVER, usando a imagem hiperespectral HYDICE, com as quatro (4) classes. De acordo com os resultados, o método MAXVER apresentou uma separação entre as classes, com uma acurácia global de 81,50% e coeficiente kappa de 0,45. Este valor de acurácia obtido ao empregar o tradicional MAXVER explica que 81,50% dos pixels na imagem foram corretamente classificados, e 18,5% dos pixels tiveram erros na classificação.

O método MAXVER foi aplicado com a imagem original, considerando todas 162 bandas após o pré-processamento, e classificou somente as feições maiores das 4 classes. Esta limitação na classificação das classes é atribuída ao facto de as feições não classificadas apresentarem um pequeno número de amostras de treinamento em um espaço de características de alta dimensionalidade, o que é confirmado na literatura por LANDGREBE (2003) e CAMPS-VALLS et al. (2014).

O MAXVER resultou numa acurácia do produtor de 100% para asfalto, 0,46% para grama e 0% para ambas classes árvore e telhado. Os erros por inclusão ao classificar as classes asfalto, árvore, grama e telhado foram respetivamente 0,10%, 100%, 0% e 0%. Durante o processo de classificação, as classes asfalto e árvore não sofreram omissão, contrariamente as classes grama e telhado que sofreram omissão em respetivamente 99,54% e 100%. As acurácias do usuário para classes asfalto, árvore, grama e telhado foram respetivamente 99,90%, 0%, 100% e 0%.

O valor do coeficiente kappa para o experimento usando o MAXVER foi menor que 0,9 (0,45), o que de acordo com Cohen (1960) e Landis e Koch (1977), sugere uma fraca concordância entre as classes.

4.2.5 Análise Conjunta dos Resultados Obtidos pelos Métodos I – IV, Usando a Imagem Hiperespectral HYDICE



TABELA 13 - RESUMO DA ACURÁCIA DOS EXPERIMENTOS I – IV COM IMAGEM HYDICE.

	PAC-MAXVER				PA-MAXVER				PCA-MAXVER				MAXVER			
Acurácia (%)	APr	EO	AU	EI	APr	EO	AU	EI	APr	EO	AU	EI	APr	EO	AU	EI
Asfalto	97.5	2.5	99.7	0.4	95.6	4.4	99.3	0.7	97	3	99.9	0.1	100	0	99.9	0.1
Árvore	99.4	0.6	85.1	14.9	98.6	1.4	86.7	13.3	98.9	1.1	89	11	0	0	0	100
Grama	92.6	7.4	98.6	1.5	92.1	7.9	96.5	3.5	93.9	6.1	96.9	3.1	0.5	99.5	100	0
Telhado	97.6	2.4	98.1	1.9	97.1	2.9	89.9	10.1	99.1	0.9	95.7	4.3	0	100	0	0
AG	96.47				95.26				96.69				81.50			
CK	0.95				0.93				0.95				0.45			

FONTE: O AUTOR, 2018.

Os experimentos realizados usando a imagem hiperespectral HYDICE estão resumidos na Figura 23 e Tabela 13. Para os experimentos foram consideradas as 4 classes presentes na imagem. Foi proposto o método de classificação PAC-MAXVER, e para efeitos comparativos, o PA-MAXVER, PCA-MAXVER e MAXVER.

De acordo com os resultados apresentados, a classificação usando o método PAC-MAXVER resultou numa acurácia global de 96,47% e coeficiente kappa de 0,95, estando ligeiramente igualado ao PCA-MAXVER apesar de ter sido o segundo melhor método a separar as classes. O método PA-MAXVER alcançou uma acurácia global de 95,26% e coeficiente kappa de 0,93. O método PCA-MAXVER teve uma ligeira melhoria ao separar as classes se comparado com o PAC-MAXVER, com uma acurácia global de 96,69% e coeficiente kappa de 0,95. O método que apresentou menor valor ao separar as classes foi o tradicional MAXVER, com uma acurácia global de apenas 81,50% e coeficiente kappa de 0,45.

Os resultados mostram também que os valores das acurácias obtidas ao empregar estes métodos explicam que 96,47%, 95,26%, 96,69% e 81,50% dos pixels na imagem foram corretamente classificados, e respectivamente 3,53%, 4,74%, 3,31%, e 18,5% dos pixels foram classificados com erros.

O método MAXVER foi aplicado com a imagem original, considerando todas 162 bandas após pré-processamento, e classificou somente as feições maiores das 4 classes, contrariamente ao PAC-MAXVER, PA-MAXVER e PCA-MAXVER que classificaram completamente as 4 classes. Esta limitação na classificação das classes da imagem é atribuída ao facto de as feições não classificadas apresentarem um pequeno número de amostras de treinamento em um espaço de características de alta dimensionalidade, estando de acordo com LANDGREBE (2003) e CAMPS-VALLS et al. (2014).

O PAC-MAXVER resultou em uma acurácia do produtor de 97,54% para asfalto, 99,38% para árvore, 92,60% para grama e 97,56% para telhado, o que significa que os erros de inclusão nas classes asfalto, grama e telhado foram a baixo de 2% e somente a classe árvore apresentou um erro de inclusão por volta dos 15%. Durante o processo de classificação, a classe asfalto foi omitida em 2,46%, árvore foi omitida em 0,62% e atribuída por inclusão em 14,94%, grama omitida em 7,40% e telhado em 2,44%. Das classes com erros na acurácia do produtor, as classes árvore e telhado com a registo da menor acurácia no MAXVER, sendo 0% para ambas. Nos restantes métodos tais como o PAC-MAXVER, PA-MAXVER e PCA-MAXVER as mesmas classes obtiveram acurácias acima de 90%.

As melhores acurácias na classificação foram obtidas pelos métodos PCA-MAXVER e pelos métodos baseados na Propagação de Afinidade (PAC-MAXVER e PA-MAXVER), o que atestou a evidência teórica de que esses métodos tal como os restantes métodos discriminativos, não se limitam pela redução do tamanho das amostras (PENG et al., 2016). No que se refere à acurácia do produtor obtida no método PCA-MAXVER, foi em 97,04%, 98,94%, 93,92% e 99,07% para respectivamente as classes asfalto, árvore, grama e telhado. As quatro classes foram omitidas em respectivamente 2,96%, 1,06%, 6,08% e 0,93%. Os erros de inclusão neste método foram de 0,13%, 11,03%, 3,14% e 4,26% respectivamente para as classes asfalto, árvore, grama e telhado.

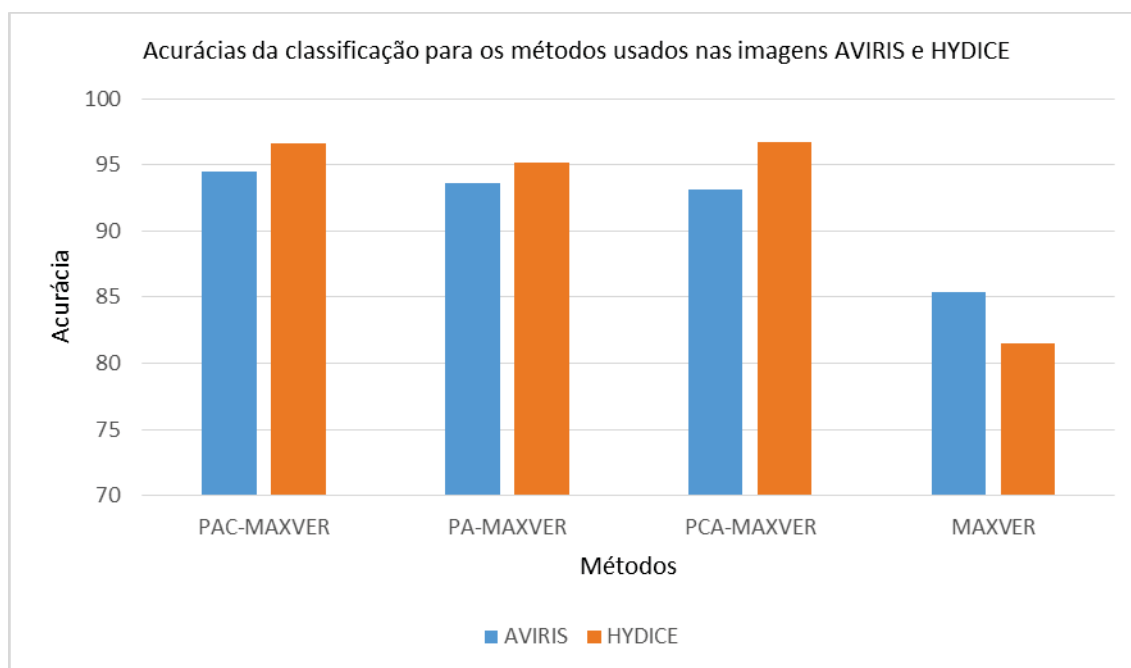
O método MAXVER, somente classificou corretamente a classe asfalto, estando as restantes 3 classes (árvore, grama e telhado) com erros de classificação. Durante a classificação, as classes foram omitidas da seguinte forma: grama (99,54%) e telhado (100%). Neste mesmo processo, foram verificados erros de inclusão, onde foram atribuídas erradamente classes para asfalto em 0,10% e árvore em 100%.

Os valores dos coeficientes kappa para os Experimentos I, II, III e IV foram respectivamente 0,95, 0,93, 0,95 e 0,45. De acordo com Cohen (1960) e Landis e Koch (1977), esses valores sugerem que os métodos PAC-MAXVER, PA-MAXVER e PCA-MAXVER que possuem o coeficiente Kappa acima de 0,9 apresentam uma concordância quase perfeita entre as classes das imagens avaliadas e boa acurácia. O método MAXVER com kappa igual a 0,45 apresenta uma fraca concordância entre as classes.

4.3 ANÁLISE CONJUNTA DOS RESULTADOS OBTIDOS NAS DUAS IMAGENS

O gráfico da Figura 24 abaixo mostra as acurácias das classificações obtidas ao aplicar os métodos PAC-MAXVER, PA-MAXVER, PCA-MAXVER e MAXVER sobre as imagens AVIRIS e HYDICE.

FIGURA 24 - COMPARAÇÃO DAS ACURÁCIAS DAS CLASSIFICAÇÕES NOS EXPERIMENTOS I – IV PARA AS IMAGENS AVIRIS E HYDICE.



FONTE: O AUTOR, 2018.

Com base nos resultados dos experimentos empregando as duas imagens, segundo o gráfico da Figura 24 e com auxílio das Tabelas 8 e 13, pode-se observar o seguinte: usando a imagem AVIRIS, os métodos propostos PAC-MAXVER e PA-MAXVER apresentaram os melhores resultados em termos de acurácia da classificação e coeficiente kappa, se comparados com PCA-MAXVER e MAXVER. Quanto aos resultados obtidos usando a imagem do sensor HYDICE que cobre uma zona urbana, os métodos testados também foram melhores, sendo superados sutilmente pelo PCA-MAXVER em apenas 0,22%. Estes resultados permitem-nos responder ao problema de pesquisa e verificar a validade da hipótese colocada nesta tese.

O problema de pesquisa desta tese foi: “Em que medida a integração do algoritmo de *clusterização* Propagação de Afinidade baseado nas Classes com o classificador paramétrico Máxima Verossimilhança melhora a acurácia da

classificação em imagens hiperespectrais”? Para responder a essa questão, é necessário que a hipótese defendida nesta tese seja verificada. Esta tese defende a hipótese de que “a integração do algoritmo de *clusterização* Propagação de Afinidade baseado nas Classes e o classificador paramétrico Máxima Verossimilhança permite classificar imagens hiperespectrais utilizando pequenas amostras de treinamento, e melhora a acurácia da classificação”.

Com base nos resultados, pode-se por um lado, observar que a classificação pelo método proposto apresentou em todos os casos (com as duas imagens) valores de acurácia superiores que os obtidos pela Máxima Verossimilhança, segundo é resumido no gráfico da Figura 24. Na imagem AVIRIS, foi obtido 94,15% para PAC-MAXVER, contra 85,42% do MAXVER. Na imagem HYDICE, PAC-MAXVER obteve 96,47% e MAXVER resultou em 81,50%. Estes resultados confirmam que o método proposto melhora a acurácia da classificação, satisfazendo parte da hipótese colocada.

Por outro lado, verifica-se que para a imagem AVIRIS o método proposto PAC-MAXVER classificou as 16 classes, mesmo nos casos em que o tamanho de amostras de treinamento era limitado. Pode-se verificar também que o método MAXVER classificou somente 10 das 16 classes existentes. Devido ao tamanho reduzido das amostras, as classes alfafa, milho-plantio-direto, grama-pasto cortado, aveia, trigo e torres não foram classificadas ao usar o MAXVER. Na segunda imagem, a do sensor HYDICE, o número de classes classificadas por ambos os métodos não alterou, isto porque para as quatro classes na imagem é possível coletar amostras maiores de treinamento. Contudo, o MAXVER não conseguiu classificar feições de tamanho menor, o que resultou numa classificação com falta de detalhes se comparada com a do método PAC-MAXVER proposto. Esses resultados confirmam a segunda parte da hipótese, segundo a qual a integração da PAC e MAXVER permite classificar imagens hiperespectrais utilizando pequenas amostras de treinamento.

Verificada a hipótese colocada nesta tese, quantitativamente, o método proposto melhorou a acurácia da classificação em imagens hiperespectrais se comparado com o MAXVER em 8,73% para a cobertura rural (imagem AVIRIS), e em 14,97% para ambiente urbano (caso da imagem HYDICE).

CAPÍTULO 5. CONCLUSÕES E TRABALHOS FUTUROS

5.1 CONCLUSÕES

Nesta tese foi proposta uma nova abordagem visando a integração do algoritmo de *clusterização* Propagação de Afinidade baseado nas Classes com o classificador paramétrico Máxima Verossimilhança para classificação das imagens hiperespectrais AVIRIS e HYDICE. Os objetivos desta pesquisa, apresentados na introdução, foram alcançados e suportados pelos resultados, chegando as seguintes conclusões:

O primeiro objetivo específico desta tese foi: extrair/selecionar características em imagens hiperespectrais através da Propagação de Afinidade, e avaliar as vantagens e desvantagens da classificação. Este objetivo foi alcançado; as características (neste caso bandas das imagens) foram extraídas/selecionadas com sucesso, confirmando os relatos da literatura sobre o potencial da PA na *clusterização*.

O segundo objetivo desta tese visava aplicar o classificador paramétrico Máxima Verossimilhança na classificação de coberturas de terra em imagens hiperespectrais, e avaliar seu potencial na classificação. De acordo com os resultados da pesquisa, este objetivo foi alcançado, a classificação foi realizada. As acurácias globais para as duas imagens foram de apenas 85,42% e 81,50% e coeficientes kappa de 0,83 e 0,45%, respectivamente para as imagens AVIRIS e HYDICE. Na primeira imagem foram consideradas 190 bandas, mas classificou somente 10 classes das 16 existentes. Para a segunda imagem foram consideradas 162 bandas e classificadas as 4 classes existentes mas com falta de algum detalhe nas classes. Esta limitação na classificação é atribuída ao fato de as classes não classificadas apresentarem pequeno número de amostras de treinamento em um espaço de características de alta dimensionalidade, estando de acordo com evidências da literatura (LANDGREBE, 2003; CAMPS-VALLS et al. (2014).

O terceiro objetivo da tese visava classificar as imagens hiperespectrais, com base na integração da PAC e MAXVER desenvolvida, e analisar o potencial da

exploração conjunta dessas abordagens. De acordo com os resultados, este método apresentou a melhor separação entre as classes, com acurácia global de 94,15% e coeficiente kappa de 0,94 para a primeira imagem e acurácia de 96,47% e coeficiente kappa de 0,95 para a segunda imagem. Esses valores do coeficiente kappa são maiores que 90%, o que de acordo com Cohen (1960) e Landis e Koch (1977), representam uma concordância na classificação e acurácia quase perfeita, o que provou a eficiência da utilização conjunta dessas duas abordagens para classificação de coberturas de terra em imagens hiperespectrais.

O quarto objetivo desta tese visava avaliar a nova abordagem de classificação desenvolvida como resultado da integração dos conceitos de *clusterização* e de classificação. A metodologia desenvolvida mostrou ser econômica e robusta sobre a acurácia da classificação, sem contudo aplicar grandes amostras de treinamento, e os resultados obtidos, de 94,15% e 96,47% respectivamente para as imagens AVIRIS e HYDICE, mostraram claramente o alcance desse objetivo. Trata-se de um método robusto que pode ser transferível para outras imagens cobrindo zonas rurais e urbanas

O quinto e último objetivo desta tese visava efetuar um estudo comparativo entre a abordagem desenvolvida e outras técnicas tais como Propagação de Afinidade original, Análise de Componentes Principais integradas à Máxima Verossimilhança e a Máxima Verossimilhança exclusiva, utilizando dados hiperespectrais dos sensores AVIRIS e HYDICE. Os resultados obtidos permitem-nos concluir que esse objetivo foi alcançado com sucesso; o método proposto PAC-MAXVER apresentou a melhor classificação com uma acurácia global de 94,15% e coeficiente kappa de 0,94. A classificação pelo método PCA-MAXVER resultou numa acurácia global de 93,24% e coeficiente kappa de 0,93, sendo o segundo melhor método a separar as classes, usando a primeira imagem. O método MAXVER apresentou a menor separação entre as classes, com acurácias globais de apenas 85,42% e 81,50% e coeficientes kappa de 0,83 e 0,45, respectivamente para as imagens AVIRIS e HYDICE. Considerando a segunda imagem, o método PCA-MAXVER apresentou melhoria sutil na acurácia da classificação (96,69%) se comparado com o método proposto PAC-MAXVER, com 96,47% de acurácia global, e ambos com coeficiente kappa de 0,95.

A metodologia proposta nesta tese mostrou-se eficiente para classificação de coberturas de terra em imagens hiperespectrais. A acurácia obtida usando o método proposto pela integração da PAC e MAXVER esteve a um nível ótimo para uso em aplicações rurais e urbanas.

No que se refere ao problema de pesquisa, pode-se concluir, com base nos resultados já discutidos, que o método proposto nesta tese melhorou a acurácia da classificação se comparado com o tradicional MAXVER em 8,73% e 8,19% para a cobertura rural (imagem AVIRIS), e em 14,97% e 13,76% para ambiente urbano (caso do sensor HYDICE).

A hipótese colocada nesta tese também foi confirmada, com base nos resultados obtidos. As acurácias obtidas foram superiores que as dos outros métodos, e a classificação detalhada até em casos de uso de pequenas amostras de treinamento tais como alfafa, milho, grama/pasto-cortado e aveia, permitindo que todas feições na imagem fossem corretamente classificadas.

O algoritmo proposto visou integrar a Propagação de Afinidade baseada nas Classes e o classificador paramétrico Máxima Verossimilhança para classificação de feições em imagens hiperespectrais. Os resultados obtidos mostraram a robustez desse método, ao classificar cenas diversificadas nas imagens de forma acurada, eficiente e rápida. Esta metodologia explorou de forma combinada as potencialidades dos dois métodos, dando um contributo na melhoria do processo de classificação.

5.2 TRABALHOS FUTUROS

Esta seção da tese visa fornecer um conjunto de recomendações de uso (práticas melhores) para a nova técnica de classificação desenvolvida. O método desenvolvido nesta tese mostrou ser econômico e robusto sobre a acurácia da classificação, usando pequenas amostras de treinamento. Porém, para que seu funcionamento seja totalmente generalizado, o método ainda carece de algumas melhorias. Tendo em vista essa consideração, sugere-se que futuros desenvolvimentos desta pesquisa estejam focalizados na integração de todas as

operações para um sistema, tornando o algoritmo automatizado simultaneamente. Sugere-se também que sejam aplicadas outras métricas de similaridade, que incluam as componentes espacial e espectral no algoritmo PAC para provavelmente melhorar ainda mais os resultados da *clusterização*.

REFERÊNCIAS

ABLIN, R.; SULOCHANA, C.H. A Survey of Hyperspectral Image Classification in Remote Sensing. **International Journal of Advanced Research in Computer and Communication Engineering**, v. 2, n. 8, 2013.

ANDREW, Y. Ng.; MICHAEL, I.J. On Discriminative vs. Generative Classifiers: A comparison of logistic regression and naïve Bayes. **14th International Conference on Neural Information Processing Systems (NIPS)**, p. 841 – 848, 2001.

ARYA, R.; ANTONISAMY, B.; KUMAR, S. Sample Size Estimation in Prevalence Studies, **Indian Journal of Pediatric**, v. 79, n. 11, p. 1482 – 1488, 2012.

BANDOS, T.V; BRUZZONE, L.; CAMPS-VALLS, G. Classification of Hyperspectral Images With Regularized Linear Discriminant Analysis. **IEEE Transactions on Geoscience and Remote Sensing**, v. 47, n. 3, 2009.

BARTELS, M., WEI, H. **Rule-based Improvement of Maximum Likelihood Classified LIDAR Data fused with coregisterd band**. Computational Vision Group, School of Systems Engineering the University of Reading, 2006.

BAUMGARDNER, M.; BIEHL, L.; LANDGREBE, D. 220 band aviris hyperspectral image data set: June 12, 1992 indian pine test site 3. **Purdue University Research Repository. doi**, v. 10, p. R7RX991C, 2015.

BENGIO, Y.; COURVILLE, A.; VINCENT, P. Representation Learning: A Review and New Perspectives. **Pattern Analysis and Machine Intelligence**, v. 35, n. 8, p. 1798-1828, 2013.

BRZANK, A.; HEIPKE, C. **Classification of LIDAR into Water and Land in Coastal Areas**. Institute of Photogrammetry and Geoinformation, 2007.

CAMPBELL, J.B. **Introduction to Remote Sensing**, Third Edition, Taylor&Francis, London, 2002.

CAMPBELL, J.B.; **Introduction to Remote Sensing**. 4th edn. New York: The Guilford Press, 2007.

CAMPS-VALLS, G.; BRUZZONE, L. Kernel-Based Methods for Hyperspectral Image Classification. **IEEE Transactions on geoscience and remote sensing**, v. 43, n. 6, June 2005.

CAMPS-VALLS, G.; TUIA, D.; BRUZZONE, L.; BENEDIKTSSON, J.A. ADVANCES IN HYPERSPECTRAL IMAGE CLASSIFICATION. **IEEE Signal Processing. Mag**, v. 31, p. 45 - 54, 2014.

CENTENO, J.A.S. **Sensoriamento Remoto e Processamento de Imagens Digitais**, UFPR, Editora CPGCG, ISBN 85-88783-05-03, 2009.

CHANG, C.I. **Hyperspectral imaging: techniques for spectral detection and classification**, Kluwer Academic, New York. (Ed.) 2003.

CHANG, C.-I.; CHIANG, S.-S. Anomaly detection and classification for hyperspectral imagery. **IEEE Trans. Geosci. Rem. Sens.**, v. 40, n. 2, p. 1314 - 1325, 2002.

CHEHDI, K.; SOLTANI, M.; CARIOU, C. Pixel classification of large-size hyperspectral images by affinity propagation. **Journal of Applied Remote Sensing**. v. 8, 2014.

CHRISTOPHER, C.; JAMES, N. **Oxford Concise Dictionary of Mathematics**. 5th edn. Oxford University Press, 2014.

CHUTIA, D.; BHATTACHARYYA, D.K.; SARMA, K.K.; KALITA, R.; SUDHAKAR, S. Hyperspectral Remote Sensing Classifications: A Perspective Survey. **Transactions in GIS**. v. 20, n. 4, p. 463 – 490, 2016.

COHEN, J. A coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, v. 20, p. 37-46, 1960.

CONGALTON, R.G. A review of assessing the accuracy classifications of remotely sensed data. **Remote Sensing Environment**. v. 37, p. 35 - 46, 1991.

COSTA, G. A. O. P.; PINHO, C. M. D.; FEITOSA, R. Q.; ALMEIDA, C. M.; KUX, H. J. H.; FONSECA, L. M. G.; OLIVEIRA, D. A. B. INTERIMAGE: an open source platform for automatic image interpretation. In: **Simpósio Brasileiro de Geomática e V Colóquio Brasileiro de Ciências Geodésicas**, 2, UNESP – Presidente Prudente. Anais - Presidente Prudente: UNESP, 2007. 2ª ed. p. 735 - 739. 2007.

CRISÓGONO, P. **Detecção remota em meio urbano: teste com classificadores supervisionados a nível do pixel sobre imagens de alta resolução espacial aplicados numa lógica hierárquica de classes**. Dissertação de Mestrado em Gestão do Território, Especialização em Detecção Remota, 2011.

CRÓSTA, A. P. **Processamento Digital de Imagens de Sensoriamento Remoto**, Campinas, P, UNICAMP, ed. rev., 1993.

DEL FRATE, F.; PACIFICI, F.; SCHIAVON, G.; SOLIMINI, C. Use of Neural Networks for Automatic Classification From High-Resolution Images. **IEEE Transactions on geoscience and remote sensing**, v. 45, n. 4, April 2007.

DUDA, R.O., HART, P.E., STORK, D.G. **Pattern Classification**. John Wiley and Son, New York, NY. p. 177, Second Edition, 2000.

DUECK, D. **Affinity Propagation: clustering data by passing messages**. University of Toronto, September 24, 2009.

ENVI, User's guide. **Research Systems Inc.**, USA, 2006.

FOODY, G.M. Status of Land cover classification accuracy assessment. **Remote Sensing of Environment**, vol. 80, p. 185 - 201, 2002.

FREY, B.J.; DUECK, D. Mixture modeling by affinity propagation. **Advances in neural information processing systems**, v. 18, p. 379, 2006.

FREY, B.J; DUECK, D. Clustering by passing messages between data points. **Science**, v. 315, n. 5814, p. 972 - 976, 2007.

FUKUNAGA, K. **Introduction to Statistical Pattern Recognition**. Academic Press Professional, Inc., San Diego, CA, USA, 1990.

GOEL, P. K., PRASHER, S. O., PATEL, R. M., LANDRY, J. A., BONNEL, R.B., VIAU, A. A. Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn. **Computers and Electronics in Agriculture**, v. 39, n. 2, p. 67 - 93, 2003.

GOETZ, A.F.H.; VANE, G.; SOLOMON, J.E.; ROCK, B.N. **Imaging spectrometry for earth remote sensing science**, v. 228, p.1147 - 1153, 1985.

GOLUBIEWSKI, N.E.; WESSMAN, C.A. Discriminating urban vegetation from a metropolitan matrix through partial unmixing with hyperspectral AVIRIS data. **Canadian Journal of Remote Sensing**, v. 36, n. 3, p. 261 - 275, 2010.

GONZALES, R.C., WOODS, R.E. **Processamento de imagens digitais**, Edgard Blucher Ltda, 1ª Edição, São Paulo, 2000.

GUAN, R, SHI, X, MARCHESE, M, YANG, C, LIANG, Y. Text clustering with seeds affinity propagation, **IEEE Transactions on Knowledge and Data Engineering**, v. 23, n. 4, p. 627 - 637, 2011.

HOFFBECK J.P.; LANDGREBE D.A. Covariance Matrix Estimation and Classification with Limited Training Data. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v.18, n. 7, July 1996.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. **Journal of Educational Psychology**, v. 24, p. 417 - 441, e 498-520, 1933.

HSU, P.-H. Feature extraction of hyperspectral images using wavelet and matching pursuit. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 62, n. 2, p. 78 – 92, 2007.

HUANG, C.; DAVIS, L.S.; TOWNSHEND, J.R.G.; An Assessment of Support Vector Machines for Land Cover Classification. **International Journal of Remote Sensing**, v. 23, n. 4, p. 725 - 749, 2002.

JENSEN, J.R. **Introductory Digital Image Processing: A Remote Sensing Perspective**, Pearson Prentice Hall, New Jersey, USA, 1996.

JIA, S.; JI, Z.; QIAN, Y.; SHEN, L. Unsupervised Band Selection for Hyperspectral Imagery Classification Without Manual Band Removal. **IEEE Journal of selected topics in applied earth observations and remote sensing**, v. 5, n. 2, April 2012.

JIA, S.; QIAN, Y. “Constrained nonnegative matrix factorization for hyperspectral unmixing,” **IEEE Transactions on Geoscience and Remote Sensing (TGRS)**, v. 47, n. 1, p. 161 - 173, 2009.

JIA, S.; QIAN, Y.; JI, Z. Band Selection for Hyperspectral Imagery Using Affinity Propagation. **Digital Image Computing: Techniques and Applications, IEEE Computer Society**, DOI 10.1109/DICTA.2008.42, 2008.

JIA, X. Simplified maximum likelihood classification for hyperspectral data in cluster space. **IEEE International Geoscience and Remote Sensing Symposium, (IGARSS'02)**, v. 5, p. 2578 - 2580, 2002.

JIMENEZ, L.O; LANDGREBE, D.A. Supervised classification in high-dimensional space: geometrical,statistical, and asymptotical properties of multivariate data. **IEEE Transactions on Systems, Man and Cybernetics**, v. 28, n. 1, p. 39 - 54, 1997.

JIMENEZ, L.O.; LANDGREBE, D.A.; Hyperspectral Data Analysis and Supervised Feature Reduction Via Projection Pursuit. **IEEE Transactions on Geoscience and Remote Sensing**, v. 37, n. 6, p. 2653 – 2667, 1999.

LANDGREBE, D.A. Hyperspectral image data analysis. **IEEE Signal Process. Mag.** v. 19, n. 1, p. 17 - 28, 2002.

LANDGREBE, D.A. **Signal Theory Methods in Multispectral Remote Sensing**, 1st ed. Hoboken, NJ: Wiley, 2003.

LANDIS, J.R.; KOCH, G.G. The measurement of observer agreement for category data, **Biometrics**, v. 33, n. 1, p. 159 - 174, 1977.

LEE, C.; LANDGREBE, D. A. Analyzing high-dimensional multispectral data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 31, p. 792 - 800, 1993.

LEVESQUE, J., STAENZ, K. Monitoring mine tailings revegetation using multitemporal hyperspectral image data. **Canadian Journal of Remote Sensing**, v. 34, p. S172 - S186. doi: 10.5589/m07-068, 2008.

LI, J.; QIAN, Y. Dimension Reduction of Hyperspectral Images with Sparse Linear Discriminant Analysis. In **IEEE International Geoscience and Remote Sensing Symposium (IGARSS)**, p. 2927 - 2930, July 2011.

LILLESAND, T.M.; KIEFER, R.W.; CHIPMAN, J.W. **Remote Sensing and Image Interpretation**, John Wiley & Sons, Hoboken, NJ, USA, 2004.

LIU, X.; XIA, W.; WANG, B.; ZHANG, L. "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," **IEEE Transactions on Geoscience and Remote Sensing (TGRS)**, v. 49, n. 2, p. 757 - 772, 2011.

LONG, Y.; LI, HENG-CHAO, CELIK, T.; LONGBOTHAM, N.; EMERY, W.J. Pairwise-Distance-Analysis-Driven Dimensionality Reduction Model with Double Mappings for Hyperspectral Image Visualization. **Remote Sensing**, v. 7, p. 7785 - 7808, 2015.

MANOLAKIS, D.; MARDEN, D.; SHAW, G.A. Hyperspectral image processing for automatic target detection applications, **Lincoln Laboratory Journal**, v. 14, p. 79 - 116, 2003.

MARANGOZ, A. M., KARAKIŞ, S., ORUÇ, M., SAHIN, H., SEFERCIK, U. G., TOPAN H., BÜYÜKSALIH, G. 3D Cultural Heritage Documentation of Safranbolu Test Site Using High Resolution Satellite Imagery, **1st Workshop of the EARSeL Special Interest Group Urban Remote Sensing**, Berlin, Germany, 2-3 March 2006.

MASELLI, F.; CONESE, C.; PETKOV, L.; RESTI, R. Inclusion of prior probabilities derived from a nonparametric process into the maximum likelihood classifier. **Photogrammetric Engineering and Remote Sensing**, v. 58, p. 201 - 207, 1992.

MATHER, P.M. **Computer Processing of Remotely Sensed Images: an Introduction**. St Edmundsburi Press Ltd, 1987.

MEILA, M.; SHI, J. Learning segmentation by random walks. In Leen, T.K.; Dietterich, T.G.; Tresp, V, **Advances in Neural Information Processing System**, v. 13, p. 873 - 879, 2001.

MENESES, P.R.; DE ALMEIDA, T.; ROSA, A.N.C.S.; SANO, E.E.; DE SOUZA, E.B.; BAPTISTA, G.M.M.; BRITES, R.S. **Introdução ao processamento de imagens de sensoriamento remoto**. Brasília 2012.

MULLER, K.-R.; MIKA, S.; RATSCH, G.; TSUDA, K.; SCHOLKOPF, B. An introduction to kernel-based " learning algorithms. **IEEE Neural Networks**, v. 12, n. 2, p. 181 - 201, 2001.

NG, A.Y.; JORDAN, M.I.; WEISS, Y. On spectral clustering: analysis and an algorithm. In. Dietterich, T.G.; Becker, S.; Ghahramani, Z., editors, **Advances in Neural Information Processing Systems**, v. 14, p. 849 - 856, 2002.

NG, A. Y.; JORDAN, M. I. On Discriminative vs. Generative classifiers: A comparison of logistic regression and naïve Bayes. **NIPS**, 2002.

PAL, M.; MATHER, P.M. Some issues in the classification of DAIS hyperspectral data. **International Journal of Remote Sensing**, v. 27, p. 2895 – 2916, 2006.

PANDYA, M.; BAXI, A.; POTDAR, M.B.; KALUBARME, M.H.; AGARWAL, B. Comparison of Various Classification Techniques for Satellite Data. **International Journal of Scientific & Engineering Research**, v. 4, n. 1, p. 1 - 31, 2013.

PEARSON, K. On Lines and Planes of Closest Fit to Systems of Points in Space. **Philosophical Magazine**. v. 2, n. 11, p. 559 - 572, 1901.

PEDDLE, D.R.; BOULTON, R.B.; PILGER, N.; BERGERON, M.; HOLLINGER, A. Hyperspectral detection of chemical vegetation stress: evaluation for the Canadian HERO satellite mission, **Canadian Journal of Remote Sensing**, v. 34, n. 1, p. S198 - S216, 2008.

PENG, Y.; PAN, Z.; ZHENG, Z.; LI, X. Hyperspectral Image Classification by Fusion of Multiple Classifiers. **International Journal of Database and Theory and Application**, v. 9, n. 2, 2016.

PERREIRA, J.L.G. **Estudos de Áreas de Florestas em Regeneração através de Imagens Landsat TM**. Dissertação de Mestrado. São José dos Campos: INPE, 1996.

QIAN, Y.; JIA, S.; ZHOU, J.; ROBLES-KELLY, A. "Hyperspectral unmixing via sparsity-constrained nonnegative matrix factorization," **IEEE Transactions on Geoscience and Remote Sensing (TGRS)**, v. 49, n. 11, p. 4282 - 4297, 2011.

QIAN, Y.; YAO, F.; JIA, S. Band selection for hyperspectral imagery using affinity propagation, **IET Computer Vision**, v. 3, n. 4, p. 213 – 222, 2009.

RAMAKRISHNAN, D.; BHARTI, R. Hyperspectral remote sensing and geological Applications. **Current Science**, v. 108, n. 5, 10 March 2015.

REFIANTI, R.; MUTIARA, A.B.; SYAMSUDDUHA, A. A. Performance Evaluation of Affinity Propagation Approaches on Data Clustering. **(IJACSA) International Journal of Advanced Computer Science and Applications**, v. 7, n. 3, 2016.

RIBEIRO, S., SCALEA, R., CONEGLIAN, F., OLIVEIRA, P. 2009. Classificação supervisionada de imagens orbitais de diferentes resoluções radiométricas. **Anais XIV Simpósio Brasileiro de Sensoriamento Remoto**, INPE, p. 1013 - 1020.

RICHARDS, J.A. **Remote Sensing Digital Image Analysis**. Springer-Verlag: Berlin, Germany, p. 281, 1986.

RICHARDS, J.A. **Remote Sensing Digital Image Analysis: An introduction**. Springer-Verlag: Berlin, Germany, p. 240, 1999.

RICHARDS, J.A.; JIA, X. **Remote Sensing Digital Image Analysis: An Introduction**. Berlin, Germany: Springer, 2006.

RICHARDS J.A. Supervised Classification Techniques. **In: Remote Sensing Digital Image Analysis**. Springer, Berlin, Heidelberg, DOI https://doi.org/10.1007/978-3-642-30062-2_8, 2013.

ROESSNER, S., SEGL, K.; HEIDEN, U., KAUFMANN, H. Automated differentiation of urban surfaces based on airborne hyperspectral imagery. **IEEE Transactions on Geosciences and Remote Sensing**, v. 39, n. 7, p. 1525 - 1532, 2001.

ROSALES, R; FREY, B.J. Learning generative models of affinity matrices. **Proc UAI**, 2003.

SACHIN, S. Overcoming variations in nanometer-scale technologies. **IEEE Journal on Emerging and Selected Topics in Circuits and Systems**. v. 20, n. 1, p. 5 - 18, 2011.

SAHOO, B.C.; OOMMEN, T.; MIRSA, D.; NEWBY, G. Using the one-dimensional S-transform as a discrimination tool in classification of hyperspectral images. **Canadian Journal of Remote Sensing**, v. 33, n. 6, p. 551 – 560, 2007.

SAHOO, S.R.; PANDA, P.K.; RAY, P. K. Hyperspectral image analysis for iron ore discrimination in Konjhar district, Odisha. **International Journal of Remote Sensing & Geoscience (IJRSG)**, v. 4, n. 2, March 2015.

SCEPAN, J. Thematic validation of high-resolution global land-cover data sets, **Photogrammetric Engineering and Remote Sensing**, v. 65, p. 1051 - 1060, 1999.

SERPICO, S. B; BRUZZONE, L. A. New Search Algorithm for Feature Selection in Hyperspectral Remote Sensing Images. **IEEE Transactions on Geoscience and Remote Sensing**, v. 39, n. 7, July 2001.

SHENTAL, N.; ZOMET, A.; HERTZ, T.; WEISS, Y. Pairwise clustering and graphical models **Neural Information Processing System**, n.16, 2003.

SOLOMON, C.; BRECKON, T. **Fundamentals of Digital Image Processing: A Practical Approach with examples in MATLAB**, John Wiley & Sons, Chichester, UK, 2011.

STORY, M.; CONGALTON, R. Accuracy assessment: a user's perspective, **Photogrammetric Engineering and Remote Sensing**, v. 52, p. 397 - 399, 1986.

STREHL, A.; GHOSH, J. Cluster ensembles - a knowledge reuse framework for combining partitions. **Journal of Machine Learning Research**. v. 3, p. 583 - 617, 2002.

SULSOFT (2003). Aster DTM 2.0 Installation and User's Guide. **Technical report**, Sulsoft Ltd, Porto Alegre, Brazil, 2003.

TARABALKA, Y.; CHANUSSOT, J.; BENEDIKTSSON, J. A. Segmentation and classification of hyperspectral images using watershed transformation. **Pattern Recognition**, v. 43, n. 7, p. 2367 - 2379, 2010.

VAIPHASA, C. Innovative genetic algorithm for hyperspectral image classification. **Proceedings of MAP ASIA 2003 conference**, Kuala Lumpur, Malaysia, 2003.

WEILER, N. Honeypots for Distributed Denial of Service Attacks. **Eleventh IEEE International Workshops on Enabling Technologies**, 2002.

WEN, X.P.; YANG, X.F.; HU, G.D. **Hyperspectral remote sensing data mining using multiple classifiers combination**. In Ponce, J and Karahoca, A (eds) Data Mining and Knowledge Discovery in Real Life Applications. Vienna, Austria, I-Tech, 2009.

YANG, C.; BRUZZONE, L.; SUN, F.; LU, L.; GUAN, R.; LIANG, Y. A Fuzzy-Statistics-Based Affinity Propagation Technique for Clustering in Multispectral Images. **IEEE Transactions on geoscience and remote sensing**, v. 48, n. 6, June 2010.

YANG, C.; BRUZZONE, L.; GUAN, R.; LU, L.; LIANG, Y. Incremental and Decremental Affinity Propagation for Semisupervised Clustering in Multispectral Images. **IEEE Transactions on geoscience and remote sensing**, v. 51, n. 3, March 2013.

ZHANG, C.; QIU, F. Mapping individual tree species in an urban forest using airborne lidar and hyperspectral imagery. **Photogrammetric Engineering & Remote Sensing**, v. 78, n. 10, p. 1079 - 1087, 2012.

ZHANG, X, WU, F, XIA, D, ZHUANG, Y, Partition Affinity Propagation for Clustering Large Scale of Data in Digital Library. **College of Computer Science**, Zhejiang University, Hangzhou, China, 2008.

ZHU, Y; HUANG, C. An Improved Median Filtering Algorithm for Image Noise Reduction. **Physics Procedia**, v. 25, p. 609 - 616, 2012.